

Peace and honour

The 2005 Nobel Peace Prize is a timely reminder of the good work done by the International Atomic Energy Agency and its director, Mohamed ElBaradei.

Recipients of the Nobel Peace Prize have not always met with universal approval, but the 2005 award to the International Atomic Energy Agency (IAEA) and its director, Mohamed ElBaradei, is both timely and appropriate.

Since it was set up within the United Nations in 1957, the IAEA has sought to slow the spread of nuclear weapons while allowing nations access to the peaceful use of nuclear technology. In the past few years, it has played an important role in verifying adherence to the Nuclear Non-Proliferation Treaty and restricting the spread of weapons technology. But with greater support from governments, it could do far more.

At a time when nuclear non-proliferation faces several acute challenges, the IAEA has managed to maintain its reputation as an impartial arbiter on nuclear issues. This is due in large part to the performance of ElBaradei, an Egyptian-born lawyer and diplomat, who has fought fiercely to maintain the agency's independence. One of his finest hours was in 2001, when he refused to confirm the Bush administration's contention that the Iraqi government had restarted its nuclear programme. In a testament to his diplomatic acumen, ElBaradei survived a subsequent US attempt to oust him from his directorship, and was reappointed to a third term by unanimous consent in June.

But the IAEA could do a lot more to confront nuclear threats if the national governments whose contributions keep it going allowed it funding and authority equal to its task.

First and foremost, the IAEA should be given wider powers to enforce the non-proliferation treaty (see *Nature* 435, 132; 2005). Under a proposed modification to the treaty, the IAEA's powers to inspect facilities would be strengthened and its remit to analyse nations' nuclear intentions widened. Most parties to the treaty have

implemented the additional protocol, but others — including the United States, Russia and Iran — have yet to do so.

Second, the agency should be given more support to help it secure nuclear facilities and recover missing nuclear materials. At present, the agency's directorate for safety and security, which oversees such activities, has a budget of just US\$22.4 million per year — a meagre amount for such a complex and important task. Given the frequency of nuclear materials trafficking, nations should be prepared to sharply increase this budget. The agency is not meant to be a nuclear police force, but more funding would allow it to help member states ensure the security of their own nuclear materials, while continuing recovery operations in the countries in the former Soviet Union, which are riddled with abandoned sources.

Finally, nations should give serious consideration to ElBaradei's proposal that the world's nuclear-fuel stockpiles be put under the stewardship of his agency. Such a proposal may sound radical, but if implemented carefully, it could go a long way towards ensuring that most of the world's nuclear material is accounted for. More importantly, placing the distribution of nuclear fuel

"Nations should give serious consideration to ElBaradei's proposal that the world's nuclear-fuel stockpiles be put under the stewardship of his agency."

under an international authority would discredit the claims of countries such as Iran and Brazil that they must develop dual-use, nuclear fuel technology to meet their domestic energy needs.

All these steps will require the commitment of the IAEA's member states. Unfortunately, the most important players are unable to agree on how to move the agency forward. The award of this year's peace prize should serve as a wake-up call for them. ■

From rhetoric to reality

President Bush's acknowledgement of the threat of pandemic flu is welcome, if belated.

Researchers have long warned of the threat of an avian flu pandemic. The message has been taken on board by some politicians, notably in Canada, Britain and Australia. But the past few weeks have seen an unprecedented flurry of top-level US diplomatic activity about flu, prompted in large part by recognition of the fact that the public would not tolerate an inadequate response to another major catastrophe after Hurricane Katrina.

Last month, George Bush told the United Nations that the United States would be at the centre of an international coalition to fight the threat of a pandemic. Last week, he convened officials from more

than 80 countries to take the plan forward, and then sat down with the heads of the vaccine industry. The Senate has approved \$3.9 billion to fight avian flu, with \$3 billion of it going on a US stockpile of antivirals. And Michael Leavitt, the health secretary, has gone on a ten-day trip to meet officials in affected countries in southeast Asia.

But behind this activity, there are disconcerting gaps between the official discourse and reality. US officials have correctly identified promptly stamping out outbreaks as a top priority, emphasizing the need for affected countries to share data and samples. This is an important issue, but what can the coalition do if China and Vietnam, for example, refuse to share data? And even more urgently required are funds to build surveillance capacity in vulnerable countries, and to eradicate the disease in livestock.

The United Nations' Food and Agriculture Organization estimated in February that a minimum of \$100 million is needed to begin tackling the problem effectively. But so far, a few countries



(including the United States) have pledged a total of just \$16.5 million. The outbreak teams of the World Health Organization (WHO) are dwarfed by the challenge facing them.

Many scientists in affected countries are reluctant to cooperate with what often seems to them a one-way street. Hospitals in these countries often have barely enough antivirals to treat existing cases, and lack diagnostics. A better atmosphere for sharing will only come if rich nations offer these countries true cooperation and substantial aid in research and health infrastructure to deal with outbreaks.

US officials have in the past weeks made strident warnings about the lack of preparedness for a pandemic, while simultaneously giving overly reassuring messages that the job is now well in hand. Much of the \$3.9 billion in the Senate bill would go towards buying drugs, for example. But US officials have neglected to mention that the United States currently has only enough drugs ordered to cover 1% of its population. Ten countries already have drugs ordered to cover 25–40% of their populations, but it will be several years before the United States can match this, as it is at the back of the queue at the door of Roche, the sole supplier of Tamiflu.

Roche's monopoly on the drug, and its inability to ramp up production swiftly to meet demand, are themselves cause for concern. UN secretary-general Kofi Annan last week hinted that countries might use compulsory licences to produce the drug off-patent, arguing that intellectual-property rights should not be allowed to get in the way of access by the poor to medication. This option has been

rejected by the WHO, which recently obtained a donation of 3 million courses from Roche.

With few drugs on the horizon, US officials are stressing that the key weapon in the event of a pandemic will be a vaccine, and that the biggest bottleneck is in industry's capacity to produce it in sufficient quantities. But the vaccine in question is a prototype. It requires such huge doses that, even if the entire world vaccine production capacity of 900 million doses of seasonal flu vaccine antigen were switched to making it, just 75 million doses could be manufactured. In contrast, antigen-sparing strategies using adjuvants could allow from 1 billion to 7 billion doses to be produced, and this is why all efforts should be directed to this goal.

Although research into new vaccines and drugs may help us fight pandemics better several years down the line, the frontline of a pandemic would be in communities and hospitals, so this is where governments need to focus their disaster planning. For example, taps and door knobs in washrooms are a significant route of flu transmission. Converting them to be pushed or opened with an elbow, as in surgical areas of hospitals, could cut transmission during a pandemic.

Vaccines and antiviral drugs deserve top-level attention. But so too do much simpler means of protecting citizens. ■

"The frontline would be in communities and hospitals, so this is where governments need to focus their disaster planning."

Advise the president

The merger of two White House advisory panels sends out the wrong message.

The US President's Council of Advisors on Science and Technology (PCAST) has rarely fulfilled the full potential of its nominal role, which is to provide the most powerful elected official in the world with scientific advice.

In theory, the presidentially appointed panel could keep the president informed on key science- and technology-related issues, ranging from avian flu and global warming to computer viruses and nuclear-weapons proliferation.

In practice, however, the panel has never lived up to that ideal. It came closest, perhaps, under the first President Bush, who graced PCAST meetings with his presence. The panel was active but not particularly influential under Bill Clinton, and has been almost invisible under the current president.

So the news that PCAST is to be merged with another, even more obscure panel, the President's Information Technology Advisory Committee (PITAC), will make few waves. Nonetheless, the amalgamation of the panels, and the expansion of the possible number of members from 25 to an unwieldy 45, portends a possible weakening of the voice of science in the White House.

PCAST has already confined itself to the relative arcana of science policy. At the moment, for example, it is evaluating the effectiveness of the National Nanotechnology Initiative — a worthwhile exercise, but hardly one that is likely to grab the president's attention.

Floyd Kivimäki, a venture capitalist who co-chairs PCAST with John Marburger, the president's science adviser, says the new panel will operate much as before, with the new work delegated to appropriate subcommittees. But unless the panel becomes considerably more active, its new role overseeing all the information-technology research initiatives in the federal government may mean that less time and resources are available to work on science issues. This marks a continuation of the tendency of the Bush administration to marginalize the voices of science in its internal deliberations.

One of the difficulties that will always face a body such as PCAST is the sheer vastness of the territory it is supposed to cover. These days, advice on specific scientific questions will often require detailed specializations that few PCAST members will possess. At the same time, there is a tendency for officially designated advisory bodies that are required by US law to meet in public — such as PCAST — to shun robust discussion of substantive issues. Finally, the president's discretion in appointing the entire panel himself is not conducive to the delivery of solid and occasionally unwelcome advice.

PCAST is the latest in a series of similar panels stretching back to the administration of Harry Truman. Some have been more active and influential than others, depending largely on the president's own interest in science and his relationship with the chief science adviser. Perhaps a future administration will develop the committee's role and profile instead of neglecting it — but even then, the panel's pre-eminence will last only as long as that president's term in office. ■

"The president's discretion in appointing the panel himself is not conducive to the delivery of solid or unwelcome advice."

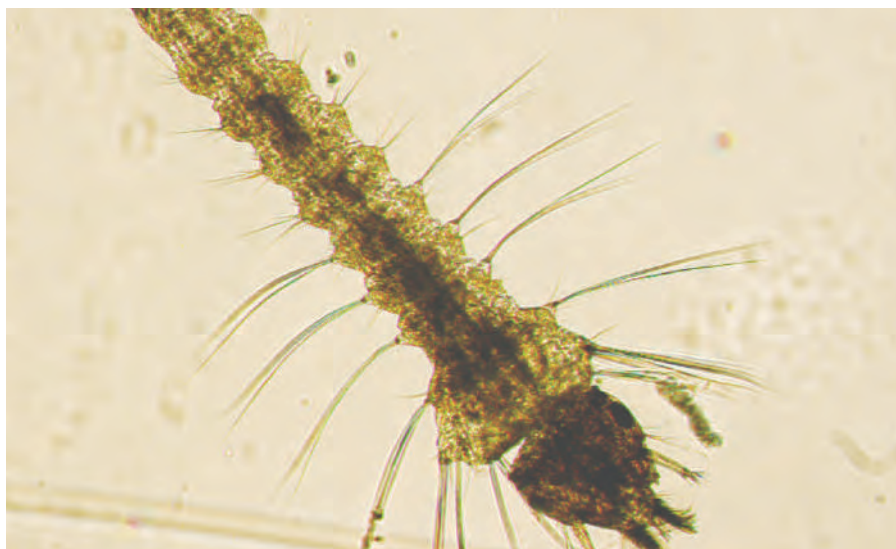
RESEARCH HIGHLIGHTS

Glimmer boys

Nature Biotechnol. doi:10.1038/nbt1152 (2005)

Biologists have genetically engineered mosquitoes to produce males with sperm that glows. The point? The fluorescent gonads make it possible to separate males from females at an early stage of larval development.

Insect control strategies often depend on the release of sterile insects. But in mosquitoes, females are the transmitters of malaria, so it is important to release only males. The glowing genitalia offer a way to isolate a male-only population that can then be sterilized and let out into the wild. The researchers at Imperial College London, led by Andrea Crisanti, carried out their experiments on the mosquito *Anopheles stephensi* (larva pictured right), which is the principal vector of human malaria in Asia.



F. CATERUCCIA & A. CRISANTI

CELL BIOLOGY**Lipids link to Alzheimer's**

Nature Cell Biol. doi:10.1038/ncb1313 (2005)

Elevated levels of the amyloid beta peptide ($A\beta$) are known to lead to the formation of brain plaques (pictured below, damaged areas in brown) in Alzheimer's disease. But the normal function of $A\beta$ has been a mystery. Now a study reveals a role for it in regulating lipid metabolism.

Tobias Hartmann of the Centre for Molecular Biology in Heidelberg, Germany, and his colleagues engineered cells from mice to lack the machinery that makes $A\beta$. These cells ended up containing large amounts of cholesterol and lipids.

The researchers also identify $A\beta$'s targets: two enzymes that are both involved in

cholesterol and lipid-processing. The work should help to explain the association that has been noticed between Alzheimer's and high cholesterol levels.

MOLECULAR ELECTRONICS**Rattling chains**

Phys. Rev. B **72**, 121405(R) (2005)

The way molecular vibrations affect the flow of electrons through a chain of atoms has been captured in a new charge transport model.

Developed by researchers from Ohio University in the United States and from the Pontifical Catholic University in Rio de Janeiro, Brazil, the model takes into account the repulsive effects between electrons, and how the molecule's vibrations, called phonons, affect the energy levels that the electrons can occupy. This makes it more realistic than previous models, which focused on only one of these two interactions.

The researchers found that vibrations can open unexpected transport channels, which allow electrons to hop between atoms, or they can block the electrons' flow. These observations may help in modelling the behaviour of electrical components made from single molecules.

ANTIBIOTIC RESISTANCE**Short enzyme stymies drug**

J. Biol. Chem. doi: 10.1074/jbc.M505727200 (2005)

The troublesome resistance of members of the *Mycobacterium tuberculosis* complex to a new class of macrolide antibiotics called ketolides stems from a truncated form of a gene called *erm*.

Like other *erm* genes, which are found in a diverse range of pathogenic bacteria, the *erm*(37) homologue in *M. tuberculosis* encodes a methyltransferase enzyme with a highly targeted action. By adding a methyl group to a particular nucleotide in a specific ribosomal RNA sequence, this enzyme confers resistance to older macrolide antibiotics, such as erythromycin, but not to the newer ketolides.

Researchers led by Stephen Douthwaite of the University of Southern Denmark, Odense, now show that ketolide resistance arises from the imprecise action of *erm*(37)'s enzyme. The enzyme encoded by the truncated gene slips along its RNA target, adding extra methyl groups to neighbouring nucleotides.

IMMUNOLOGY**Neurons enter the fray**

J. Virol. **79**, 12893-12904 (2005)

Human neurons are capable of sensing and responding to viral infection, finds a group led by Monique Lafon at the Pasteur Institute in Paris.

Previously, it was thought that neurons relied on infection being detected by their companion glial cells. But Lafon's group discovered that infecting neurons with rabies virus; a herpes virus; or with double-stranded RNA, a molecular signature of RNA viruses, switched on genes involved in immunity.

Neurons were also found to express the protein Toll-like receptor 3, which recognizes and responds to double-stranded RNA. Double-stranded RNA could therefore be the trigger that turns on the innate immune response in human neurons.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

CELL BIOLOGY

Comfortably numb

J. Neurosci. **25**, 8924–8937 (2005)

Camphor, the pungent active ingredient of mothballs, produces a warming sensation and mild local anaesthesia when applied to the skin. But although camphor has been used therapeutically for centuries, its mode of action has only just been uncovered.

Using rat neurons, David Clapham and his colleagues at the Howard Hughes Medical Institute in Boston showed that camphor works in the same way as capsaicin. This is the 'hot' compound in chilli peppers, and also a mild analgesic. Both substances activate, and subsequently numb, the heat-sensitive TRPV1 receptor of sensory neurons. Camphor was also found to inhibit the cold-sensitive TRPA1 receptor.

MATERIAL SCIENCE

Bonds writ large

Appl. Phys. Lett. **87**, 131903 (2005)

Even big cracks start small: the collapse of a bridge begins with the breaking of atomic bonds at the tip of a flaw. But studying real materials at an atomic scale during fracture is very challenging. So Francisco Emmerich of the Federal University of Espirito Santo in Brazil is offering an alternative. He has designed a scaled-up solid, using bar magnets stacked in a brick-wall arrangement and separated by layers of foam.

The model solid closely mimics the forces between atoms. Experiments using it show that catastrophic failure always starts in the same way: two of the magnets at the crack tip jump apart to a critical separation. This may be equivalent to a chemical bond breaking at the atomic scale.

EVOLUTIONARY GENETICS

Copy and save

Genome Res. **15**, 1421–1430 (2005)

A key question in genetics is how extra copies of genes manage to persist in an organism. When a gene becomes duplicated, one copy often gets deleted or inactivated. But sometimes both remain intact.

To find out more, Uwe Sauer and his team at the Swiss Federal Institute of Technology in Zurich used computer modelling and quantitative biochemical experiments to study 105 families of duplicated genes in the yeast *Saccharomyces cerevisiae*.

They found that spares adopt four survival strategies: a copy can act as a backup, develop a new function, become regulated differently

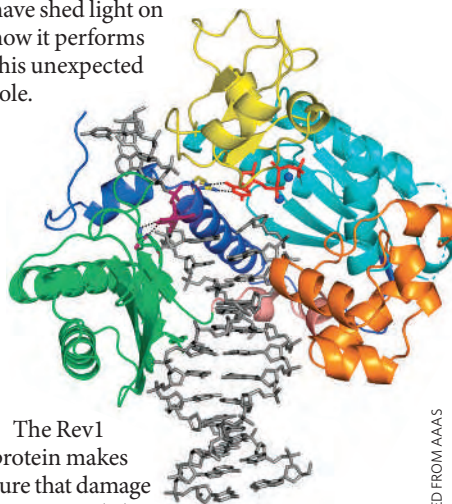
or simply boost the levels of the protein encoded by that gene. All mechanisms seem to be equally important.

PROTEIN STRUCTURE

Skipping the G spot

Science **309**, 2219–2222 (2005)

The protein pictured below has a remarkable ability: it can stand in for a genetic base when DNA is being copied. By deciphering the crystal structure of the protein, Aneel Aggarwal of the Mount Sinai School of Medicine in New York and his colleagues, have shed light on how it performs this unexpected role.



The Rev1 protein makes sure that damage to guanine (G) bases in a DNA template can be bypassed. Without the protein, a spoilt base would block gene replication. Aggarwal's structure shows yeast Rev1 bound to a template G (violet) in a DNA helix (grey). The protein evicts the base from the helix and attracts the G's partner, a cytosine (C) base, to bind to itself. This means the C, carried by the molecule dCTP (red), gets incorporated into the new DNA strand in the right place.

ADAPTED FROM AAAS

CANCER

Extra control on suicide

Cell **123**, 49–63 (2005)

The p53 gene controls cell suicide and division, helping to stop cells becoming cancerous. After DNA damage, levels of the p53 protein increase. This happens because the protein is destroyed less quickly, but also, says Michael Kastan, because p53 translation increases. Kastan and his colleagues at the St Jude Children's Research Hospital in Memphis, Tennessee, found that ribosomal protein L26 binds to p53 messenger RNA and enhances its translation into protein. They also found a second protein, nucleolin, which inhibits p53 translation.

JOURNAL CLUB

Bruno Sicardy

Paris Observatory, and Pierre and Marie Curie University, Paris

An astronomer tells of the missing link in Saturn's rings.

The number of satellites known to be roaming between Saturn's rings and its moon Enceladus has soared over the past few decades as telescopes and instruments have improved. With the arrival of spacecraft, such as the current orbiter Cassini, satellites are even being discovered very near or inside the rings themselves.

Recently, Carolyn Porco of the Space Science Institute, Boulder, and her colleagues in the Cassini imaging team revealed six new moons in the ring region — typically no more than 5 km across (C. C. Porco *et al. Science* **307**, 1226–1236; 2005). Since this paper was published, a seventh moon has been detected in the narrow Keeler gap, which lies inside the A ring.

These new moons may hint at a missing link between the collisional, fluid-like rings and fully formed satellites. Some of them have a surprisingly low density, and their discovery has blurred the distinction between what is a satellite and what is merely a clumpy aggregate of dust.

Porco *et al.* estimated the density of some other moonlets by looking at the ripples they excite in the rings. For instance, Atlas, which orbits just 900 km outside the main rings, has a density of 0.5 g cm^{-3} . This is in the same ballpark as the values my colleagues and I obtained using Earth-based observations for the familiar moons, Prometheus and Pandora (S. Renner *et al. Icarus* **174**, 230–240; 2005).

The region just outside the main rings of Saturn is acting as a natural laboratory where we can admire short-lived, loose and fluffy aggregates as they emerge from their native rings. I expect that a continuous collision and re-accretion process transforms rings into satellites and vice versa with turnover times of a few tens of millions of years.

NEWS

US progressives fight for a voice in bioethics

WASHINGTON DC

Conservative bioethicists often provide intellectual ammunition for US politicians on major issues ranging from stem-cell research to right-to-die decisions. Now several prominent researchers are joining forces to promote different scientific values in public debate.

Arguing that conservative bioethics is out of step with most Americans, the group is forming a 'progressive' movement to influence discussions of scientific and medical topics. "It is important for progressive bioethics to enter the political fray," says Arthur Caplan, an ethicist at the University of Pennsylvania, Philadelphia.

On 3 October, members of the group set out key elements of their approach at a meeting at the Center for American Progress, a left-leaning think-tank in Washington DC that is helping to start the movement. They define themselves in part by what they oppose: the conservative stance embraced by Republican political leaders, by right-leaning think-tanks, and by the President's Council on Bioethics under Leon Kass, who led the council until 1 October (see *Nature* 437, 307; 2005).

But the progressive group hopes to emulate the conservatives' success in influencing public policy. Conservative bioethicists have set up a network of think-tanks and journals that issue position papers, book media appearances and hold meetings with politicians.

These strategies have shaped the Republican response in debates over stem-cell research and the right-to-die case of Florida's Terri Schiavo. Caplan and others were outraged when Republican leaders fought to keep Schiavo on life support against her husband's wishes. "Nothing could make clearer the difference between progressive and conservative bioethics," says Caplan.

Ethicists at the meeting say that their approach is optimistic about science and technology. "Progressive bioethics opens itself to change," says Alta Charo of the University of Wisconsin, Madison. The conservative approach, they argue, often focuses on how technology could adversely affect the essence of humanity. In a letter accompanying a 2002 report from the President's Council on Bioethics, for example, Kass told the US

president that human cloning "carries with it a number of troubling consequences for children, family, and society".

The progressive bioethicists say they plan to study topics not often covered by conservatives — such as inequities in the healthcare system. These inequities were highlighted by Hurricane Katrina, which left thousands of poor African-Americans stranded without federal assistance for almost a week. "Progressive bioethics has to talk about those who are left behind and left out," says Vanessa Gamble of Tuskegee University in Alabama.

"Progressive bioethics has to talk about those who are left out."

And the group hopes to avoid the political missteps that have sometimes resulted from conservative approaches. Public opinion polls found that the Republican efforts to keep Schiavo alive were unpopular. The Schiavo debate may have influenced the Center for American Progress to become more involved in ethics, adds Jonathan Moreno, who was on sabbatical there at the time. "I think they saw that it was useful to have someone like me around to put a different frame on these issues than was being set out by the conservative media," says Moreno, now a fellow at the centre.

Leaders in the progressive-bioethics movement welcome both Democrats and Republicans, saying they think ethical issues should remain bipartisan. But the group's members have supported positions taken by many Democratic politicians. And the president of the Center for American Progress, John

The Terri Schiavo case polarized opinions on life support and brought bioethics to the forefront of politics.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Nuclear group nabs peace prize

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Nobel leader —
Mohamed ElBaradei

Of all the Nobel prizes, the award for peace is the most political. After France resumed nuclear testing in June 1995, the Nobel committee awarded that year's prize to nuclear disarmament campaigner Joseph Rotblat and his creation, the Pugwash Conferences on Science and World Affairs.

The 2005 choice, say nuclear-policy experts, is equally entrenched in politics.

The winners, announced on 7 October, are the International Atomic Energy Agency (IAEA) and its director, Mohamed ElBaradei. Although the agency is tasked with impartiality while monitoring the potential spread of nuclear weapons, it has been at the centre of political disputes involving Iraq and Iran. Observers say that the prize is a signal from the Nobel committee that the agency has maintained neutrality during these rows.

Before the Iraq war, IAEA staff said there was no evidence for an ongoing nuclear-weapons programme in the country, a view vindicated by coalition inspectors after the war began. The agency is asking for more

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

C. O'NEARA/AP

Podesta, is a key Democratic strategist who served as chief of staff when Bill Clinton was president of the United States.

"There's a need to re-establish the scientific voice as a voice of fact and reason in the public dialogue," claims Podesta.

But even he isn't sure whether bioethical issues are important enough to sway the votes of Americans. That, he says, "is a political question that will work itself out over the next couple of years".

Erika Check

time to continue its work in Iran, but fears that its inspections may be curtailed because of US demands to refer Iran to the United Nations security council.

"This year's prize is clearly intended as a signal of support for multilateral diplomacy and inspections, rather than the use of military force," says Rebecca Johnson, director of the Acronym Institute for Disarmament Diplomacy in London.

The award was also a major boost for IAEA scientists who inspect nuclear facilities. Spokeswoman Melissa Fleming says that staff at the agency's Vienna headquarters were shocked and then jubilant on hearing of the award. ElBaradei had taken the day off, but hurried into the office after hearing of his prize on television news.

Is the award likely to result in anything more permanent than a glow of pride? Fleming is cautious, but says that

the prize will at least make it harder for countries to ignore ElBaradei's pleas for more funding. The agency has what ElBaradei calls a "shoestring" inspection budget of US\$100 million a year.

Others suggest that the award could strengthen the IAEA's position in arguments about whether inspections, or tougher measures such as sanctions or military force, are the best way to deal with countries with alleged nuclear-weapons ambitions. But that is wishful thinking given the current US government's antipathy towards ElBaradei, says Michael Levi, an arms-control expert at King's College London. "The award is not going to change the credibility of the IAEA in the United States," he says. "The people who don't like the agency don't like the Nobel Peace Prize."

Jim Giles

See Editorial, page 927.

ON THE RECORD

"We're uplifted. But they told us there were going to be free drinks, and there aren't any."

A member of staff at the International Atomic Energy Agency reacts after the organization wins the Nobel Peace Prize.

"My discoveries are 40 years old, and I am an old man."

Chemist Yves Chauvin describes why he felt "embarrassment, not joy" on winning the Nobel prize.

Source: *Nature*, *Der Spiegel*

SCORECARD



Adult stem cells

The Catholic Church in South Korea has found a way to avoid controversy over embryonic stem-cell research. It is planning to spend US\$10 million on adult stem-cell studies.



Bird flu

A San Diego entrepreneur may ruffle feathers with a line of avian flu-themed clothing. Among the offerings: a 'Bird Flu Tour' T-shirt and baseball caps sporting the logo 'Pandemic Fever — Catch It!'

OVERHYPED

Abstinence-only education

US health officials in the Bush administration say avoiding sex is good, but Representative Henry Waxman (Democrat, California) isn't convinced by their reasoning. Last week, Waxman charged that the National Abstinence Clearinghouse — the main group set up to evaluate abstinence-promoting programmes — is scientifically unsound. He cited several of the group's official statements, including:

"Sex therapists consider masturbation the first stage of sexual addiction for sex addicts."

"Pictures of external genitalia in any form, whether diseased or healthy, can be detrimental to the health of young men and women's minds."

The clearinghouse's stated goal is to "promote the appreciation for and practice of sexual abstinence".



Forbidden cave: without permits, the discoverers of the hobbit are unable to continue digging at Liang Bua cave.

More evidence for hobbit unearthed as diggers are refused access to cave

Even as researchers uncover more details about the ancient 'hobbit' people of Indonesia, they fear that they may never return to Liang Bua cave, where the crucial specimens were found in 2003.

"My guess is that we will not work at Liang Bua again, this year or any other year," says team leader Michael Morwood, an archaeologist at the University of New England in Armidale.

The latest findings from the cave reveal, among other details, that the metre-tall humans, known as *Homo floresiensis*, lived on the island of Flores as little as 12,000 years ago (see page 1012). But continued exploration at Liang Bua is being blocked, the researchers say, because the discovery of miniature humans runs counter to the theories of Indonesia's senior palaeoanthropologist — and national icon — Teuku Jacob.

Based at Gajah Mada University in Yogyakarta, Jacob contends that the bones are of a genetically deformed *Homo sapiens*. A condition called microcephaly, which can cause dwarfism, could explain the miniature brain, he says (see *Nature* 434, 432–434; 2005).



Stone tools could provide clues about the cultural habits of the human species discovered on Flores.

Government officials will not issue exploration permits that might prove Jacob wrong, scientists say. Neither Jacob nor the officials involved could be reached for comment.

The study of the bones that have so far been recovered has already been hindered, the researchers complain. Last winter, after holding the remains for about four months, Jacob returned them to the discovery team with some bones broken or shattered. The bones may have been damaged when casts were attempted or during transport. Today's article was delayed by at least six months, the authors say, because the bones were not available for follow-up studies. The newly described jaw, for instance, was broken in half between the front teeth, and the break has obliterated certain key structures.

"It's an outrage," says team anthropologist Peter Brown, also based at the University of New England. "It is now impossible for future scholars to verify my work. The pelvis was whole; now it is 100 crumbs."

Today's report, which details evidence of nine *H. floresiensis* individuals found at Liang Bua throughout 2004, is likely to convince any remaining sceptics that a new species of human has been identified (see page 957).

The two jaw bones found are virtually identical even though their owners lived 3,000

years apart. "You can't have a colony of microcephalics going through time," says Brown. "That's crazy."

The new bones also turned up features that are not found in modern humans. In particular, both of the jaws unearthed lack a chin structure; chins are a distinguishing feature of *H. sapiens*. The researchers also found arm bones from two individuals. "They are spectacularly long," says Brown, adding that the limb proportions are reminiscent not of *H. sapiens* but of Lucy, the 3.2-million-year-old *Australopithecus afarensis* found in Ethiopia in 1974.

But there are still many questions to answer, including how and when the species evolved, when it arrived on the island, and what its lifestyle was like (see 'The life of a hobbit'). Early this year, archaeologist Tony Djubiantono, director of the Indonesian Centre for Archaeology in Jakarta, told *Nature* that digging in Liang Bua would proceed in the summer. But he never issued the permits. Djubiantono, a co-author on today's paper, could not be reached for comment. But sources say he is reluctant to challenge Jacob and his allies in the upper echelons of the Indonesian government.

Disputes over palaeoanthropology dig sites are not uncommon — there has been considerable squabbling over the control of hominid sites in Africa. But it is unprecedented to close down such a spectacular site. "Liang Bua is the crown jewel of the caves," says Brown, adding that only a small percentage of it has been excavated so far. "This is where the team should be focusing."

Morwood says that during explorations this summer at other sites on Flores and neighbouring islands, the team has found promising hints about the origin of *H. floresiensis*, but no new hominid bones. Work in the Soa Basin, for example, suggests that hominids were present on Flores significantly earlier than 840,000 years ago, the earliest date previously reported (M. J. Morwood *et al. Nature* 392, 173–176; 1998).

"That's what you might expect in the context of some of the very primitive traits of *H. floresiensis*," says Morwood. "Historically the emphasis for early hominid studies has been Java. This may change."

But without access to Liang Bua, the mysteries of the ancient 'hobbit' people will probably remain secret for the foreseeable future. ■

Rex Dalton

The life of a hobbit

The latest archaeological finds from Flores shine further light on how life might have been for the tiny inhabitants of this Pacific outpost. The evidence suggests that *Homo floresiensis* was a decent butcher and cook, but not much of an artist.

Workers digging at Liang Bua on Flores found bones of a dwarfed form of the elephant-like species *Stegodon* alongside the tiny humans. "There are cut marks on the *Stegodon* bones, so we can infer that *H. floresiensis* was not vegetarian," says Bert Roberts, a geochronologist at the University of Wollongong, Australia, one of the latest study's authors.

Some of the bones are charred, he adds, which suggests cooking. "There are also circular clusters of burnt stones, akin to 'hearth' arrangements," he adds. "If you go to the effort of making a tidy fireplace, it's not a huge



Stone tools on Flores may have been left by *H. floresiensis* or by modern humans.

step to have a barbecue."

Other experts urge caution in inferring cultural activities such as cooking from an assortment of bones and stones. "The finds might be due to modern humans," comments Chris Stringer, who studies human evolution at the Natural History Museum in London.

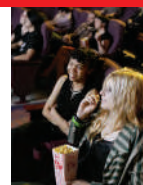
So far, no modern-human remains have been found at Liang Bua from earlier than 12,000 years ago, the time *H. floresiensis* is thought to have died out. This suggests that the two species might not have lived directly side by side. *Homo sapiens* artefacts

at Liang Bua are separated from the older material by a layer of ash from a volcanic eruption that occurred at about that time.

"But it will take only one tooth from a modern human found below the ash for the theory to be rewritten," Stringer says.

The same goes for the stone tools found at Liang Bua, says Harvard University anthropologist Daniel Lieberman. The relatively unsophisticated implements could have been made by either *H. sapiens* or *H. floresiensis*, he says. Proving that it was the latter will require uncovering a site where tools have been deliberately buried with tiny human remains.

The site contains no evidence of pigment use or bone carvings, suggesting that painting and crafts did not play a big part in the life of *H. floresiensis*. **Michael Hopkin**



SCIENTISTS DANGLE BAIT FOR SCREENWRITERS

Film summit puts a spotlight on untold stories from the lab.

www.nature.com/news

Indonesia struggles to control bird flu outbreak

As officials in Washington discuss how to tackle outbreaks of bird flu more effectively (see page 927), an outbreak in humans continues in Asia. **Declan Butler** assesses the situation in Indonesia, and finds out how likely it is that the virus might evolve into a pandemic strain.

How worrying is the Indonesian outbreak?

Previous human outbreaks have largely been in quite remote areas, but this one is centred on Jakarta, one of the world's largest conurbations and home to some 21 million people. Moreover, the country's population of 240 million looks after 1.3 billion chickens, many of which live among the 30 million backyard farms. These are spread across some 6,000 inhabited islands, making it difficult to track and control cases. Most of the cases so far have arisen on Java, home to half of Indonesia's population.

It is also unclear when or if the outbreak will end. Avian flu is now endemic in Indonesian poultry, in part because the government failed to cull birds when the virus first appeared.

What is so bad about the H5N1 strain of avian flu virus?

When scientists sequenced the genome of the flu strain that caused the devastating 1918 pandemic (see *Nature* 437, 794–795; 2005), they concluded that it probably arose from an avian flu that jumped directly to humans.

The team identified 25–30 amino-acid mutations in 1918 flu that seem to account for its virulence. Initial analysis of H5N1, which has killed 65 people in Asia so far, suggests that it already has some of these mutations. The fear is that if it accumulates more, it could acquire the ability to spread readily between humans. Each new human case gives the virus a chance to mutate and gain this ability.

How many people have the virus?

The media, the Indonesian government and the World Health Organization (WHO) all

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Under pressure: Indonesian officials are finding it hard to quantify and control the outbreak of bird flu among the country's chickens.

give conflicting numbers. This is partly because Indonesia's Ministry of Health declares cases immediately, whereas the WHO does not count cases until it has checked them in its own labs. The best current estimate, based on tests done by the Ministry of Health and the US Naval Medical Research Unit Number 2 (NAMRU-2) in Jakarta, is seven confirmed cases and two probable cases. Six of these nine individuals have died. There are also about 80 suspected cases, but most will probably prove not to be H5N1.

How well equipped is Indonesia to detect and treat cases quickly?

Outbreak investigation staff are stretched but coping, according to Steven Bjorge at the WHO's Jakarta office. Others are less sanguine. "One just has to look at the current polio, measles and dengue epidemics in Indonesia to

realize that the public-health system is having trouble coping with preventable diseases, so I don't have a lot of faith," says one scientist involved in monitoring the outbreak.

But other countries are helping out. Until 1999, Indonesia had virtually no flu-surveillance capacity. Since then, NAMRU-2, helped by the US Centers for Disease Control and Prevention in Atlanta, Georgia, has been working with the country's authorities to strengthen their ability to monitor and diagnose avian flu (C. G. Beckett *et al. Clin. Infect. Dis.* 39, 443–449; 2004). Over the past few weeks, teams have been sent to areas with poultry outbreaks to look for suspected human cases. Teams of foreign scientists, and aid, are also converging on the country.

Is the virus passing between people?

Probably, but in only a very limited way. Most cases have been spread out but there have been some family clusters. In July, an eight-year-old became sick, followed by her sister and father. On 10 September, a 37-year-old woman died in Jakarta, then her nine-year-old nephew contracted the virus. And the most recent cases — four- and five-year-old boys — are the nephews of a 20-year-old patient from Bandar Lampung in Sumatra.

But nailing down human-to-human transmission is difficult. Researchers have to work out how the first case got infected from poultry, then rule out the possibility that the other family members did not catch the virus in the same way. Most of the time, the inquiry gets nowhere. "We have no firm evidence of human-to-human transmission within a family, although sometimes it is difficult to exclude it," says Bjorge.

Has the virus mutated to make it more adapted to humans?

Not so far. "Sequences of recent human isolates are very similar to previous ones in Indonesia," says Masato Tashiro, a virologist at the National Institute of Infectious Diseases in Tokyo, which sent a team to Indonesia last week.

What are scientists worried about now?

Vietnam is bracing itself for an expected further wave of cases, as the colder temperatures during the winter months from November to February are conducive to transmission of the virus. Meanwhile, Europe is on the alert after suspected outbreaks were detected in poultry last weekend. If confirmed, they would add to fears that birds migrating from China and Russia this autumn may extend the geographical spread of H5N1. ■

A. IBRAHIM/AP

Chemical exchange captures Nobel

AP PHOTO/J. MCCOONBE

This year's Nobel Prize in Chemistry has been won by three researchers who pioneered the understanding and use of a very general class of reactions that aid the construction of carbon-based (organic) molecules.

Yves Chauvin of the French Petroleum Institute in Rueil-Malmaison showed how these 'metathesis' reactions, long familiar to the petrochemicals industry, actually work.

This understanding helped Richard Schrock at the Massachusetts Institute of Technology (MIT) in Cambridge and Robert Grubbs of the California Institute of Technology, Pasadena, to develop catalysts that made metathesis more controllable and reliable.

Thanks to the trio's work, metathesis is now widely used in the synthesis of many organic compounds, from plastics and pharmaceuticals to herbicides. "It has fundamentally transformed the way chemists think about building new molecules," says Gregory Fu at MIT, who was a postdoc with Grubbs.

The award "is what we've all been waiting for," says Steven Ley, an organic chemist at the University of Cambridge, UK. "The work is spectacularly useful and important — it is the chemistry discovery of the past 30 years."

Metathesis has taken a route — not unusual in chemistry — from industry to the academic laboratory and then back out to industry. In the process, it has been transformed from a curious observation to a useful piece of science.

"People knew there was this funny reaction

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS



Robert Grubbs (left), Richard Schrock (top) and Yves Chauvin transformed metathesis from a curiosity to a key chemical tool.

in petroleum refining," says John Brown, an organic chemist at the University of Oxford, UK. "Chauvin came up with the first serious mechanism that turned it into rational design. But it was still just a party piece until Schrock came along."

Metathesis involves two molecules that each contain a double bond between carbon atoms — the bond acts rather like a double handclasp between a pair of dancers. During the reaction, which requires a catalyst containing metal

atoms, the molecules effectively exchange partners (metathesis means 'changing places'). "The molecules are cut in half and put back together again," says Brown.

In 1971, Chauvin showed how the metal-atom catalyst coordinates this rearrangement. Soon after, Schrock began to look for more effective catalysts and, by 1990, had found that compounds containing molybdenum and tungsten were especially efficient.

But Schrock's catalysts still tended to react with other molecular groups attached to the double bonds. In 1992, Grubbs discovered that ruthenium compounds interfered less with other groups, and were more stable in air. "That really rocketed the reaction into common usage," says Ley.

"The availability of these improved catalysts has revolutionized organic synthesis," he adds. "I don't think there is a single organic chemist who hasn't used metathesis."

The award is a reminder that, in chemistry, the best contributions are often not discoveries, but tools. A general way of making new molecules can find many different applications. "Most people have these catalysts on their shelves," says Fu.

Philip Ball

AP PHOTO/C. SUZUKI

AP PHOTO/C. ENA

Ig Nobels hail world's longest-running experiment

BOSTON

Unlike other days at Harvard University, the first Thursday in October is a time when levity overtakes gravity, irreverence prevails over reason, and the flight of paper aeroplanes is actively encouraged. It is time, in other words, for the Ig Nobel Prize Ceremony.

At the fifteenth Ig Nobel ceremony, held on 6 October before a boisterous crowd in Harvard's Sanders Theatre, ten prizes were awarded to winners from a dozen countries. Winners travelled from as far as Australia and Japan to take home the 'Ig', a prize devoted to "achievements that first make people laugh, and then make them think".

The biology prize, for example, went to an international team for smelling and

cataloguing the odorous secretions of 131 species of stressed-out frog. In work that earned them the chemistry prize, two researchers from the University of Minnesota in Minneapolis showed that people can swim as fast in syrup as in water (see *Nature* doi:10.1038/news040920-2; 2004).

John Mainstone of the University of Queensland in Australia accepted the physics prize for the 'pitch-drop' experiment started back in 1927 by the prize's co-winner, the late Thomas Parnell. It shows that an ostensibly solid tar derivative can behave like a liquid, forming drops at the rate of about one every nine years.

Shortly after arriving at Queensland in 1961, Mainstone found a curious piece of equipment tucked away in a cupboard. He

had unwittingly stumbled across Parnell's experiment, by then three decades old.

Parnell, Queensland's first physics professor, had taken a sample of pitch, heated it, and placed it in a funnel. He hoped to show that this apparently solid substance — brittle enough to shatter on impact — has fluid-like properties. Sure enough, the material did form drops, albeit at an exceedingly slow rate. Its viscosity, Mainstone and his colleagues calculate, is 100 billion times that of water.

It is hard to know what motivated Parnell, but Mainstone suspects it had to do with the quantum revolution — the idea that "things are not what they seem" — that had overtaken physics. "This was his way of showing there are strange things in classical physics too," Mainstone surmises.



ROBOTIC CAR WINS GRAND CHALLENGE

Stanford team scoops \$2 million for completing cross-country course.

www.nature.com/news

STANFORD UNIV.

The experiment has attracted a cult following, he says, yet it also raises some serious scientific questions. No one knows, for example, how each drop actually detaches. Mainstone believes that fibres supporting the drop in its final stages become unstable and fail catastrophically, but this hypothesis is unconfirmed.

Part of the problem is that in the experiment's 78-year history, no one has seen a single drop fall. That's not surprising, says Mainstone. "We're talking about a descent of a few centimetres, lasting a tenth of a second, that occurs just once a decade." The last drop, which fell in November 2000, should have been recorded on a webcam, but technical problems intervened. "We'll have to wait until next time, which could be 2010 or later," Mainstone notes.

There is enough pitch left to sustain the experiment for another century, he estimates, and he hopes it will continue, despite the constant battles he has waged with the "philistines" who believe the

experiment wastes precious time and space. Mainstone's labour of love, along with Parnell's pioneering work, were recognized in 2003 when the *Guinness World Records* named the pitch-drop demonstration the world's longest-running laboratory experiment. The Ig Nobel prize, which Mainstone shares with the late

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

John Mainstone steps up to accept his award for an experiment begun almost 80 years ago.

Parnell, provides further recognition.

Looking to the future, Mainstone has already picked a successor, his former student Andrew White, to oversee the project when he finally steps down.

Mainstone never imagined that he would look after this experiment for four-and-a-half decades, but says he has become "enthralled by the historical continuity of it all". A shiny drop of pitch, which gradually changes in shape from a sphere to a pear, is "a thing of beauty", he says. While acknowledging the importance of quantum mechanics, he can't help but wonder: "Is the pitch drop any less fascinating than wave-particle duality?"

Mainstone is a great believer in the Ig Nobels, and not just because of his award. Science has become a "rat race", largely as a result of the pressure to compete for grant money, he claims, adding that it's important to get a break from that sometimes. "When we cease to see the amusing side of science, it's all over," he says. ■

Steve Nadis

REUTERS/B. SNYDER

Space agencies have bad week as craft go missing

Several space missions ran into problems last week, with Europe, Japan and Russia all suffering losses of varying degrees.

On 8 October, the European Space Agency's CryoSat mission was lost shortly after launch from northern Russia. A missing computer command caused the main engines to burn for too long, sending the probe and two of its launch stages plummeting into the Arctic Ocean. CryoSat was designed to measure the thickness of polar ice to within a few centimetres, giving scientists a three-dimensional view of ice loss at the poles.

Meanwhile, in space near the asteroid Itokawa, Japan's Hayabusa spacecraft (see *Nature* 437, 306; 2005) is hobbling on just one of three reaction wheels that control its position after a second failed on 2 October. Project engineers hope it can still collect a sample of Itokawa next month as planned.

And the Russian-European Inflatable Re-entry and Descent Technology demonstrator, an experimental capsule for returning cargo from orbit, went missing on 7 October after radio controllers lost contact during a suborbital test. As of early this week, Russian search teams had yet to locate it on the ground.

US tightens rules in battle against mad cow disease

The US Food and Drug Administration has proposed restrictions that would ban the brains and spinal cords from older cows in all animal feed. Officials say the measures will serve as better safeguards against bovine spongiform encephalopathy (BSE), also known as mad cow disease. Cattle seem to contract the disease from eating feed that contains infectious proteins, often carried by other cattle.

Critics say the planned changes would not close all the loopholes that could allow BSE

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Beefed up: US rules now prevent cow brains and spinal cord from being used in animal feed.

to spread. They say that the brains and spinal cords of cattle less than 30 months old — currently not included in the feed ban — could also harbour the infectious agent, as could cattle blood and poultry waste, which can also still enter the feed system.

A public comment period will end on 19 December, and the rules are expected to take effect next year.

Long-term child health study awaits funding

A survey of 100,000 US children is stirring to life, despite questions over its funding. The National Children's Study — a joint venture between several government agencies — plans to collect information on children from conception to the age of 21. It will look for connections between the environment and health, focusing on key health problems such as obesity, diabetes, asthma and mental health.

Six study centres, selected last month, are preparing to enrol participants. But data collection will begin only if Congress decides to fund the \$2.7-billion study — something scientists can never guarantee. "This is the way federal funding goes," says Peter Scheidt, the study's director.

Cotton pest fails to evolve resistance to transgenes

Contrary to some predictions, a study published this week suggests that a cotton pest is not evolving resistance to crops genetically modified to produce the bacterial insecticide known as *Bt*.

Building on earlier work, researchers at the University of Arizona report no evidence of resistance genes becoming more common in pink bollworm, a major cotton pest. This comes after nearly a decade of US farmers planting cotton modified with the *Bt* toxin.

Reporting online in the *Proceedings of the National Academy of Sciences* (doi:10.1073/pnas.0507857102), Bruce Tabashnik and colleagues suggest a reason why. Their mathematical model suggests that current US laws, which require farmers to plant at least 5% of their land with non-transgenic crops, are helping to delay the evolution of resistance.

Cell's decision to retract paper upsets authors

The journal *Cell* is defending its decision to retract a published paper against the authors' wishes.

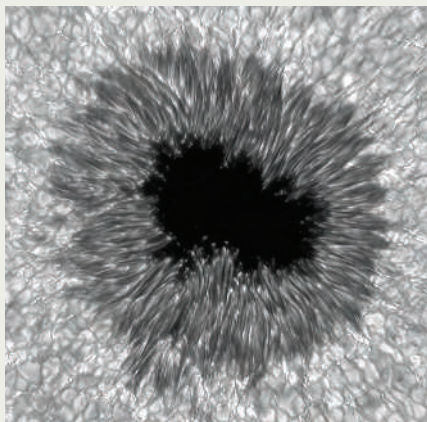
The paper in question — published in 2004 and written by Antonio Teixeira and his colleagues at the University of Brasilia — provided insights into Chagas' disease, a potentially fatal condition and a major health problem in parts of Latin America. Teixeira's team showed that the parasite that causes the disease acts by integrating its mitochondrial DNA into sufferers' genomes.

On 23 September, *Cell* announced that it was retracting the paper after becoming aware of problems with Teixeira's data on the genome site at which integration is said to occur. The decision was not endorsed by Teixeira and his colleagues, prompting several scientists to complain to *Cell* and call for the evidence behind the decision to be made public.

Last week, *Cell* editor Emilie Marcus wrote in an editorial that an outside researcher raised issues with the paper after publication. These were subsequently supported by other experts, she wrote, including one of the paper's original reviewers. Teixeira's group was offered the chance to author a retraction, but declined, she added.

Correction

Our News story on the reconstruction of the 1918 flu virus (*Nature* 437, 794–795; 2005) incorrectly located the Scientists Working Group on Biological and Chemical Weapons at the Federation of American Scientists. It is now part of the Center for Arms Control and Non-Proliferation, also in Washington DC.



Spotlight on sunspot

Sunspot AR 10810 dazzles with detail in this image, taken on 23 September by the National Solar Observatory in Sunspot, New Mexico. The sharp definition is a result of an adaptive-optics system recently installed on the 76-centimetre Dunn Solar Telescope on Sacramento Peak.

Observations of the dark spot, and the penumbral structures that radiate outward from it, can help to clarify the magnetic structure at work in sunspots. The image covers an area about 40,600 kilometres across.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Playing dirty

Forget drugs carefully designed to hit one particular molecule — a better way of treating complex diseases such as cancer may be to aim for several targets at once, says **Simon Frantz**.

It's not often that a science lecture can turn a person on to the idea of promiscuity. But when Michael Heinrich heard a talk about a promising new cancer drug, it triggered a transformation of his ideas about how to target disease. It sounds heretical, but Heinrich and others are now saying that 'magic bullet' drugs designed to hit single biological targets might not be the answer to treating complex illnesses such as cancer and cardiovascular disease. The future, they say, could be in drugs that are less picky about their molecular partners.

Heinrich's turning point was a seminar given in early 1998 at the Oregon Health and Science University Cancer Institute in Portland, where he worked. Brian Druker, a molecular biologist in the medical department at the same university, was talking about the revolutionary leukaemia treatment Gleevec (imatinib mesylate). Made by Swiss drug company Novartis, Gleevec was designed to zero in on a single protein in cancerous cells, specifically killing them while leaving healthy cells unharmed. It proved to be spectacularly effective and non-toxic. Compared with the relatively indiscriminate action and distressing side effects of conventional cancer treatments, Gleevec seemed to vindicate the single-target approach to drug discovery.

But it soon became clear that Gleevec was

not as specific as its creators had thought. The drug works by attaching to a key part of an overactive protein that causes chronic myeloid leukaemia. In his lecture, Druker revealed that the drug also inhibits a second protein, known as the PDGF receptor. Sitting in the audience, Heinrich had a brainwave. At the time, he was working on a protein similar to PDGF called KIT. "We became interested in the idea that Gleevec could probably inhibit KIT as well," he says. Working with a team led by George Demetri and Jonathan Fletcher at the Dana-Farber Cancer Institute in Boston, Massachusetts, Heinrich found that Gleevec was also remarkably effective against a rare cancer called gastrointestinal stromal tumour, known to be linked to faulty KIT activity¹.

What Heinrich and his colleagues had stumbled on went the opposite way from the direction that drug companies have been heading in since the beginning of the 1980s. In a bid to take much of the guesswork out of drug discovery, companies tried to avoid treat-

ments that non-selectively bound to several targets — what they term 'dirty' or 'promiscuous' drugs — and focused on creating selective magic bullets such as Gleevec. But researchers are now realizing that too much specificity can also be problematic.

Take aim

Before the 1980s, drug discovery began by using animal models to test compounds created by medicinal chemists. Drugs were deemed successful by virtue of their effects rather than the number of molecular targets to which they bound. For every safe and effective promiscuous drug such as aspirin, there were good treatments that caused major side effects, and plenty of other drugs that were just plain unsafe. Trying to predict side effects and understand them was almost impossible as in most cases no one knew exactly how the drugs worked.

The selective approach to drug discovery was made possible once biochemical and genetic studies began to reveal the molecular mechanisms that underlie common illnesses such as cancer and cardiovascular disease. Companies were able to pick a protein that they thought would make a good target, design compounds that interact with this protein, and test these compounds to find potential drugs.

"The idea of magic bullets is great, but in practice it's probably not going to be the right approach for complex diseases." — Bryan Roth

But 20 years down the line, it turns out that this target-based approach doesn't always guarantee success. Some of these selective drugs work in only a select population of patients. AstraZeneca's Iressa (gefitinib), for example, is designed to treat lung cancer by targeting a protein called EGFR. The drug does give an incredibly potent response, but only in about one-tenth of the patients who receive it². And, as Gleevec fortuitously showed, treatments that block more than one target can be tolerated better than previously thought.

The idea that promiscuous drugs might be more effective than targeted ones has also been emerging from efforts to understand how antipsychotic drugs work. The schizophrenia drug Clozaril (clozapine), for example, works because it targets a large number of proteins, says Bryan Roth, a biochemist at Case Western Reserve University in Cleveland in Ohio. Variations designed to bind to fewer targets and reduce Clozaril's unpleasant side effects don't work as well and still have similar side effects³.

In the 1990s, Roth and his team investigated which nerve-cell receptors were being targeted by a range of antipsychotic drugs. They found that the drugs that bound to the most receptors were the most successful in the clinic. "What became clear to us when we examined antipsychotic drugs was that the more targets they hit the better," says Roth.

Multiple choice

The reason for this is that common disorders such as cancer, cardiovascular disease and depression tend to result from multiple molecular abnormalities, not from a single defect. What's more, pinpointing a single target is unlikely to help in many cases because cells can often find ways to compensate for a protein whose activity is affected by a drug, a phenomenon known as redundancy.

Using what Roth calls 'magic shotguns' to target multiple points in these complex systems, could reap bigger therapeutic rewards than fully blocking one target. "The idea of the magic bullet continues to be a great idea, but in practice it's probably not going to be the right approach for complex diseases," says Roth.

Findings such as those of Heinrich, Roth and their colleagues have triggered a recent shift in efforts to create drugs that hit more than one target simultaneously. A number of companies and research groups are now screening compounds that stick to several targets, or are even trying to engineer promiscuous drugs. Arguably, the biggest area for promiscuous drugs at the moment is cancer⁴.

A key set of targets includes enzymes called kinases. Many of these, such as EGFR, influence how cells divide and are often abnormally active in cancers. A slew of treatments (see "Dirty" drugs under development, right) that block several kinases together are now in clinical trials. The hope is that these will work better than highly selective treatments, and that hitting more than one kinase at once will

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Hitting the spot: Gleevec was seen as proof that the 'magic bullet' approach to drugs was a success.

reduce the chance of tumours becoming resistant to the drugs. In August, Pfizer submitted a cancer drug called Sutent (sunitinib malate), acquired when it bought the biotech company SUGEN, for approval to the US Food and Drug Administration. The drug blocks not only the proteins targeted by Gleevec, but also two other similar molecules⁵.

Nevertheless, researchers in the field are frustrated that large drug companies seem to be ignoring the advantages of promiscuity, or polypharmacology as it is sometimes known. "There is still a perception in the field that multi-kinase inhibitors are going to be inherently toxic and non-selective and it's absolutely untrue," says Julie Cherrington, executive vice-president for research and development at Phenomix in San Diego, California, who helped develop Sutent when she was at SUGEN. "I can remember having conversations about this when we started to develop Sutent, and I'm still having these conversations now."

Big pharmaceutical companies are largely

'Dirty' drugs under development

Sutent

Pfizer has submitted Sutent to the US Food and Drug Administration (FDA) for approval as a therapy for kidney and gastrointestinal cancer.

Sorafenib

Created by Bayer and Onyx Pharmaceuticals, this treatment for kidney cancer is currently being considered by the FDA for approval.

Zactima

Made by AstraZeneca, Zactima is undergoing final (phase III) clinical trials in lung cancer.

AG-013736

Designed by Pfizer, this drug is undergoing efficacy (phase II) clinical trials for kidney and thyroid cancer.

still wedded to the 'one-target one-disease' model, and it's not easy to change this culture, says Simon Mencher, principal at Natrogen Therapeutics in Milwaukee, Wisconsin. "The first person I ever talked to in a large company about promiscuous drugs said: 'I agree with you but I can't convince the management to change the way they work,'" he says. "Too much funding has been sunk into targeting single agents."

Culture shock

Andrew Hopkins, head of knowledge discovery at Pfizer in Sandwich, UK, agrees that the single-target approach remains the main strategy in big companies. But this is now being challenged by fresh information on some compounds, as well as by models mimicking the effect of compounds on cells. In addition, large-scale genetic projects have confirmed the extent of redundancy by showing that altering the activity of many genes one at a time may have limited clinical effect⁶. "Polypharmacology isn't new, what is new is the realization of its importance in efficacy," says Hopkins.

But screening for compounds that hit multiple targets is a difficult task. Unlike the single-target strategy, in which the compound selected is generally the one that sticks best to the target, the most likely candidate for a multi-target drug will be one that moderately influences several targets positively and negatively at the appropriate concentrations.

Overcoming this problem requires a deeper understanding of the cellular mechanisms at which the drug is aimed. To tackle this, researchers have turned to the emerging field of network biology, which can model the complex interactions between all the molecular constituents of a cell⁷. By building these networks, researchers can identify molecules and processes that are altered in diseases. They can also predict whether it is better to design drugs that hit multiple points in one process or that dampen parallel processes, and whether redundancy will be a factor.

If multiple-kinase inhibitors prove successful in the clinic, they could drive more efforts towards promiscuous drugs. Already the Gleevec story is having an impact in industry, says Roth. He has noticed a subtle change in the drugs that large companies are licensing from smaller companies. "Both Pfizer and Merck have licensed relatively non-selective antipsychotic compounds," says Roth. "The fact that they are doing this shows that they're getting the message."

Simon Frantz is news editor for Nature Reviews Drug Discovery.

- Joensuu, H. *et al.* *N. Engl. J. Med.* **344**, 1052-1056 (2001).
- Cohen, M. H. *et al.* *Clin. Cancer Res.* **10**, 1212-1218 (2004).
- Roth, B. L., Sheffler, D. J. & Kroeze, W. K. *Nature Rev. Drug Discov.* **3**, 353-359 (2004).
- Mencher, S. K. & Wang, L. G. *BMC Clin. Pharmacol.* **5**, 3 (2005).
- Mendel, D. B. *et al.* *Clin. Cancer Res.* **9**, 327-337 (2003).
- Zambrowicz, B. P. & Sands, A. T. *Nature Rev. Drug Discov.* **2**, 38-51 (2003).
- Barabási, A.-L. & Oltvai, Z. N. *Nature Rev. Genet.* **5**, 101-113 (2004).

The forgotten ecosystem

Everyone knows about the Amazon rainforest, but Brazil's tropical savannah is arguably under greater threat. **Emma Marris** visits a testing ground for future conservation strategies.

On the savannah of central Brazil, it is dry and hot, but far from lifeless. Iridescent blue butterflies the size of handkerchiefs fly by. Leafcutter ants march in line across the dusty red soil, holding snippets of foliage like sails. The silica-filled leaves of shoulder-high shrubs clatter like snare drums in the breeze.

Yet just over the next rise lies another landscape: the unvarying green of a soya field. In a dramatic change in land use, this vast inland region of savannah and dry woodland, known as the Cerrado, is rapidly being replaced with crops and pasture. Over the past 35 years, more than half of the Cerrado's original expanse of two million square kilometres has been taken for agriculture. It is now among the world's top regions for the production of beef and soya.

Agriculture is one of the largest and most dynamic parts of Brazil's economy, and those working to save the Cerrado are unlikely to be able to slow or stop the sector's expansion. Just 2.2% of the Cerrado is protected, and it is losing ground faster than the Amazon rainforest to the north. At the current rate of loss, the ecosystem could be gone by 2030, according to estimates by Conservation International in Washington DC (see map, opposite).

Like the Amazon basin, the Cerrado is a great source of biodiversity. Its 137 threatened species include the maned wolf (*Chrysocyon brachyurus*), a striking, long-legged beast that resembles a fox on stilts. And the sparse, scrubby vegetation features more than 4,000 species that grow only here.

Yet the Cerrado has little of the global recog-

nition or celebrity advocacy that has helped advance the cause of conservation in the Amazon. If some of the savannah is to be saved, conservationists have realized, they will have to stress its importance to Brazil's economic well-being. This may also mean working hand-in-hand with those who are developing the Cerrado for agriculture — risking accusations of 'greenwash' from purists in the environmental movement. Given that similar pressures face other neglected ecosystems, the Cerrado could become an important barometer of the prospects for conservation worldwide.

In from the cold

The Cerrado might still have been an overlooked ecosystem had it not been for a project to document the workings of the world's tropical savannahs, set up in the mid-1980s by the United Nations Educational, Scientific and Cultural Organization. Under that project, a group of students at the University of Brasilia were given their start in ecology. "Seven of us went abroad to get our PhDs, and we formed a little revolution," says Carlos Klink, who is now a professor at the university. Today, they are in the vanguard of efforts to characterize and preserve what remains of the Cerrado.

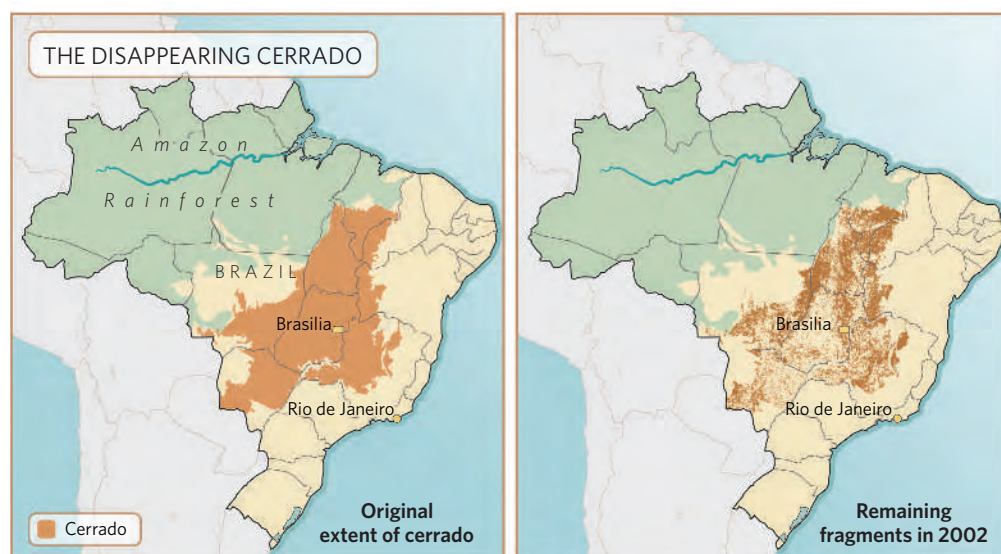
At his field site, 35 kilometres south of Brasilia, Klink shows off the pits he has dug to study the Cerrado's underground life. Below the surface, huge root systems stretch to suck up water and nutrients from the poor soil. Klink points to fields that he has burned off after various intervals to see how the natural fire cycle affects the pattern of vegetation.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The push to grow soya beans (above) is destroying the habitat of endangered species such as the maned wolf (top).

"This is a pequi," he says, grasping a skinny tree with wilted-looking leaves. Its fruits have a yellow pulp and taste divine. "It's pollinated by bats," he adds. The pequi tree (*Caryocar brasiliense*), at three metres, is the tallest thing in this bit of the Cerrado — an untidy patchwork of tall floppy grasses, greyish bromeliads and nondescript scrubs with puckered leaves.



It's certainly not the archetype of scenic beauty. "Even scientists in Brazil tend to think that this is a secondary kind of ecosystem," says Klink. "That's a pity. I find it a lovely place."

It is also a fragile ecosystem, which can be severely damaged by fires that start on neighbouring pasture. The savannah depends on natural fires, started by lightning, to clear out dead grasses and pop open seed-pods. "But you must have the right amount of fire," says Roberto Cavalcanti, another University of Brasília ecologist, who is currently working with Conservation International in Washington DC. Alas, the oily molasses grass (*Melinis minutiflora*), which was once widely planted for pasture and has since invaded the fringes of the wild Cerrado, can cause fires to rage so hotly that they burn through the tough fire-adapted bark of native woody plants.

Fires are also set deliberately to clear land for pasture. But the soil is so poor that, without extensive use of fertilizer, the imported grasses tend to fail after two or three years of busy grazing, leaving just dust, rock and carbuncular termite mounds several feet high.

Price of progress

For hundreds of years, the Cerrado's inaccessibility and poor soil saved it from large-scale exploitation. But as Brazil embraced the Green Revolution in the 1970s, new soya varieties and fertilizers made the region a viable agricultural prospect. By then, the gleaming new administrative capital of Brasília had been built in the midst of the area, bringing with it roads and people, eager for fresh economic opportunities.

For a while, clearing of savannah was encouraged by the government, because it eased development pressure on the Amazon. Today, there is official recognition of the need for conservation in the Cerrado. But still the odds remain stacked in favour of the rainforest. Brazil's Forest Code requires owners of Amazon land to set aside 80% as a reserve. In the Cerrado, the requirement is just 20%, and enforcement is poor.

Ruth DeFries, a geographer at the University

of Maryland in College Park, can testify to the problems. She is keeping track of changes in land use in Brazil using data from the Terra and Aqua satellites. But it is not easy because the Cerrado and pasture look very similar from low orbit. So she travels periodically to the region of Mato Grosso, where the Cerrado meets the rainforest, to verify her satellite observations. In Mato Grosso alone, she has seen some one million hectares of Cerrado and forest cleared for cropland between 2001 and 2004. "I have seen an awful lot of clearing," she says.

Slowing this destruction will mean working with the government and agricultural interests, say ecologists. And the window of opportunity may be small. "These are the crucial five years," says Cavalcanti. "Remember, conservation is cheaper than restoration."

John Buchanan, director of agriculture and fisheries at Conservation International's Center for Environmental Leadership in Business in Washington DC, is grasping the opportunity in partnership with Bunge, the largest producer of soya beans in Brazil, headquartered in White Plains, New York. In a pilot project, Buchanan has worked with Bunge employees and local farmers to set up the legally required reserves so that a contiguous corridor of Cerrado snakes across the landscape. "We really have no choice but to work with agribusiness," says Buchanan. "It's not an either/or situation."

Trading places

In July, when the Society for Conservation Biology held its annual meeting in Brasília, such projects came under scrutiny in a session on assessing trade-offs between human and ecological needs. Some representatives from the state-run Brazilian Agricultural Research Corporation, known as EMBRAPA, attended warily, worried that their role in bringing about the large-scale development of the Cerrado would see them cast as villains.

Geraldo Martha Júnior, an agronomist at EMBRAPA based in Brasília, gave one of the most provocative presentations, arguing that the best way to avoid unnecessary encroachment into the Cerrado is to ensure that introduced pasture grasses are adequately fertilized. He also considered the option of pasturing cattle on natural vegetation. But his presentation showed that it wasn't economically viable: the cows just wouldn't get fat enough to turn a profit.

By the end of the session, the EMBRAPA scientists seemed less cautious — and by and large, their presentations were received graciously by the assembled conservation biologists. In part, that reflects a sense of relief that research into protecting the Cerrado is now under way. Until recently, Brazil's savannah was a minority scientific interest, and there was no organized effort to save it from the plough.

The welcome extended to Martha Júnior and his colleagues is also indicative of a growing realization among conservationists that their strategies must accommodate economic development (see *Nature* 437, 614–616; 2005). To this end, ecologists working in the Cerrado are now stressing the 'ecosystem services' it provides — many of which have a tangible economic value. Some are investigating the role of the native landscape as a carbon sink, as a centre of genetic diversity for the crop cassava, or as a protector of Brazil's soil and water.

Klink is working to map the region in terms of carbon storage and movement. Native pasture, with its huge root systems, stores more carbon than pasture and crops. He and his team have also quantified the amount of water held in different ecosystems. They found that in planted pasture, evaporation equalled rainfall, whereas in the Cerrado it was about 20% less, keeping more water available in the ground.

João Campari, director of the central savannah programme at the Nature Conservancy in Brazil, has taken this approach a step further and proposed a Cerrado Grassland Exchange — a financial instrument that would attach monetary, tradable value to the areas that are protected by Brazil's Forest Code.

DeFries is saddened by the thought of more savannah being cleared for the inevitable expansion of agriculture. But pragmatism has to be the order of the day, she says. "Agricultural production and the revenue that it provides is important for the development of Brazil. I don't think it's effective to take the viewpoint that the entire Cerrado be conserved," says DeFries. "That's not realistic or desirable."

DeFries is saddened by the thought of more savannah being cleared for the inevitable expansion of agriculture. But pragmatism has to be the order of the day, she says. "Agricultural production and the revenue that it provides is important for the development of Brazil. I don't think it's effective to take the viewpoint that the entire Cerrado be conserved," says DeFries. "That's not realistic or desirable."

DeFries is saddened by the thought of more savannah being cleared for the inevitable expansion of agriculture. But pragmatism has to be the order of the day, she says. "Agricultural production and the revenue that it provides is important for the development of Brazil. I don't think it's effective to take the viewpoint that the entire Cerrado be conserved," says DeFries. "That's not realistic or desirable."

Emma Marris is a correspondent for *Nature* in Washington DC.

"We really have no choice but to work with agribusiness. It's not an either/or situation."

— John Buchanan

The maestro of minds

With a mathematician's logic and the perfectionism of a concert pianist, Nikos Logothetis is making waves in cognitive neuroscience — and putting the German town of Tübingen on the scientific map. **Alison Abbott** pays him a visit.

Nikos Logothetis could have been a professional musician: as a student, he used to play the piano in Greek bars to earn his tuition fees. And his undergraduate training in mathematics nearly took him into physics. But his bedside reading in those formative years eventually led him to biology, for which neuroscientists can count themselves lucky.

Today, Logothetis is at the top of his field. His experimental rigour commands widespread respect, and he is churning out a stream of red-hot papers. More importantly, his meticulous experiments have advanced the brain-imaging technique of functional magnetic resonance imaging (fMRI) as a respectable scientific tool.

Logothetis seems both pleased and slightly embarrassed by such accolades. During our interview, he frets about seeming conceited as we discuss his achievements. Such are the troubles of a bone fide polymath. Logothetis speaks six languages, and studied music for seven years at the Athens Conservatory in Greece. But music was always a companion, not a profession. His first love was mathematics. Yet after completing a degree in the subject from the University of Athens in 1977, Logothetis was unsure about how to apply his training.

First, he toyed with the idea of going into physics, but worried that the certainties of 'hard science' would prove insufficiently challenging. Then, soon after graduating, he fell under the spell of a book on his bedside table: *Chance and Necessity* by Jacques Monod, the French molecular geneticist who shared the 1965 Nobel prize in medicine for his work on the genetic regulation of enzyme synthesis. Logothetis was particularly struck by Monod's chapter on molecular cybernetics, which introduced him to the concept of systems control in biology. "That chapter was a true revelation for me and decisive in starting my biology studies," Logothetis says.

Universal talent

Although Logothetis had previously dismissed biology as being "too descriptive and soft", he immediately embarked on a second undergraduate degree in the subject at the Aristotle University of Thessaloniki. There, he was introduced to neurobiology, which became his professional obsession only at the very end of the course.

After rushing through a PhD in neuroscience at the Ludwig-Maximilians University in Munich, Logothetis moved to the United



The expanding Max Planck Institute for Biological Cybernetics in Tübingen is luring top neuroscientists.

States for 12 years. He first worked at the Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts, as a postdoctoral fellow and later as a research scientist, and then joined the faculty of the Division of Neuroscience at Baylor College of Medicine in Houston, Texas.

Logothetis's main research focus has always been visual perception — understanding what goes on in our brains when we perceive and recognize objects. During his early career, Logothetis worked mostly by recording the activity of single neurons using sensitive electrodes; he still uses such techniques for a major part of his work. But during the 1990s, Logothetis became increasingly involved in fMRI. This variant of MRI was then becoming a fashionable research tool, but was also acquiring a degree of notoriety. Some scientists were making exaggerated claims for its ability to locate the precise anatomical locations of particular cognitive processes.

Logothetis has no patience with such lack of rigour. He admits to having an "obsessive-compulsive attention to detail", and soon gained a reputation for careful and elegant fMRI studies. The brain is a very accommodating structure," he warns. "It will let you generate a mass of data and interpret them to support your idea." The key, he says, is strict quality of methodology and keeping the ears resolutely plugged against the siren song of over-interpretation.

More fundamentally, Logothetis was frustrated that no one had fully tested the underlying assumption behind fMRI. In most fMRI studies, the main signal recorded,

known as BOLD, is a measure of changes in oxygen levels in the brain, caused by variations in blood flow. But until Logothetis entered the fray, nobody had confirmed that these fluctuations actually correlate with neuronal activity.

Head hunting

In 2001, after painstaking studies in which he compared traditional electrophysiological recordings in anaesthetized monkeys with fMRI scans, Logothetis and his colleagues finally pinned down the neural correlate of the BOLD signal. This turned out to be 'local field potentials' — not the firing of individual cells, but the more slowly varying electrical activity of groups of cells as they receive and integrate information¹. The paper was immediately hailed as a *tour de force* of experimental dexterity and data analysis², and in 2003 was the most highly cited study in biology, apart from a handful of genome papers.

By the time his landmark paper was published, Logothetis had been back in Europe for several years. In the mid-1990s, he had been in hot demand, and was vigorously pursued by Germany's Max Planck Society, which installed him as director of the Max Planck Institute for Biological Cybernetics in Tübingen. There, he was given the best available facilities, including high-magnetic-field-strength fMRI machines, and state-of-the-art facilities for caring for experimental primates.

But happy as he was to be back in Europe, it wasn't long before Logothetis started to feel oppressed by the isolation of Tübingen — a smallish provincial town away from major



Multi-talented: now a leader in the field of cognitive neuroscience, Nikos Logothetis studied music for seven years and could have been a concert pianist.

intellectual centres. Then, in 1997, came an offer to head the prestigious new McGovern Brain Research Institute at MIT, being set up with a donation of around US\$350 million from the computer publishing magnate Pat McGovern and his wife Lore Harp McGovern.

Extraordinary steps were taken to keep Logothetis in Germany. Hubert Markl, the president of the Max Planck Society at the time, worked hard to find a way to transfer Logothetis's entire department to a larger city. Frankfurt, which boasts a high level of neuroscience research activity, was front runner.

Done deal

Negotiations opened up, but regional politics blocked the efforts. So Markl vowed instead to bring the mountain to Mohammed. A critical mass of high-level neuroscientists must be attracted to Tübingen, he declared. Markl aimed to recruit a leading technical expert in brain imaging, who could land Tübingen in pole position in the race to develop the next generation of fMRI machines. That person was Kamil Ugurbil from the University of Minnesota, who has now been appointed director of the Tübingen Institute's new department for high-field magnetic resonance. A €22-million (US\$27-million) building is currently being built to accommodate Ugurbil's team and three new fMRI machines.

Meanwhile, the German Research Foundation (DFG), Germany's main research granting agency, lobbied for a new neuroscience institute planned by the charitable Hertie Foundation to be established in Tübingen. The Hertie Institute for Clinical Brain Research

opened its new building in May last year. With 150 staff, the institute will work on conditions such as Alzheimer's disease.

Safe in the knowledge that a critical mass of neuroscientists are gathering around him, Logothetis is now settled. He has expanded his own department to more than 60 people and his head is a cauldron of plans. He is still not convinced that we understand enough about the neural events underlying fMRI recordings. The local field potentials that correlate with the BOLD signal, for instance, are affected by many different components of the brain's electrical activity. "We must understand the

"I prefer to work more like a postdoc than a research director."

— Nikos Logothetis

relation between the fMRI signal and all those components," says Logothetis.

Providing the answers will mean pushing the experimental envelope on all fronts. Logothetis plans to use 'tetrodes' — electrodes with four contact points — to follow in more detail what is going on in several neurons simultaneously. Those recordings will then be integrated with studies in neuropharmacology, optical spectroscopy and molecular imaging. Making sense of the resulting data will push even Logothetis's mathematical skills to the limit.

Logothetis is also using high-field fMRI machines as spectrometers to record the levels of individual neurotransmitters. And together with chemists at the University of

Tübingen, he is developing a sophisticated MRI contrast agent which changes its magnetic properties when it is exposed to calcium, allowing the calcium to be visualized. This, he hopes, will provide a direct signal for the fluxes of calcium ions that accompany the transmission of neural impulses. "If the molecular signalling works, then maybe we would not need to know so much about the BOLD signal," he jokes.

Meanwhile, Logothetis's meticulous fMRI studies continue to make waves among neuroscientists. In one recent paper¹, he argued that the visual cortex does not reorganize itself following damage to the retina. If Logothetis is right, he will have overturned a large body of literature which claims, using recordings from single neurons, that such reorganization does occur. Although the jury is still out, few who are familiar with Logothetis's work would bet against him being right.

Unlike many lab chiefs, Logothetis says he will continue to deploy the hands-on skills that made his name. "I prefer to work more like a postdoc than a research director," he says. "I like to do experiments myself." Typically, he runs experiments in the mornings and crunches data until the small hours. Then it's time for the maestro to wind down: if he's lucky, he fits in half an hour on his piano before a short night's sleep.

Alison Abbott is Nature's senior European correspondent.

1. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. *Nature* **412**, 150–157 (2001).
2. Raichle, M. E. *Nature* **412**, 128–130 (2001).
3. Smirnakis, S. M. *et al.* *Nature* **435**, 300–307 (2005).

BUSINESS

The technology trap

The widely admired US system for transferring ideas from the lab to the marketplace is showing signs of distress. Virginia Gewin reports.

For years, the US university innovation machine has been the envy of the world. The system — which gives universities the patent rights to technologies and lets them exploit these rights pretty much as they see fit — is generally credited with nurturing Silicon Valley and building the powerful US biotechnology industry.

Yet there are murmurs that the system isn't quite as successful as it seems. Universities' revenue from patent licences is modest: last year, it totalled \$1.3 billion, just 3% of their total research spending. Many good ideas are still left sitting on the shelf. Industrial corporations in sectors such as information technology and biotechnology say that their interactions with universities are problematic. For the first time since 1980 — when the Bayh-Dole Act handed over property rights from the federal government to the universities — there is a growing movement to overhaul the way in which technology transfer works in the United States.

"Universities are struggling in their drive to become more entrepreneurial," says Don Siegel, an economist at Rensselaer Polytechnic Institute in Troy, New York, and editor of the *Journal of Technology Transfer*. Indeed, most technology-transfer offices at universities fail to cover their own expenses, much less generate revenue streams. Siegel adds that meeting local politicians' expectations for research-related economic development is another pressure on public universities. At the same time, industry is tiring of disputes over intellectual property and, in some cases, withdrawing from collaboration with universities.

There is a widespread view in some industries that the university's technology-transfer offices are too focused on biotechnology, where a single, blockbuster patent is likely to be the key to a successful commercial product. In most other sectors, a single patent is just part of a complex mosaic of ideas that contribute to a product. Much of the drive to reform the system comes from major information-technology companies, such as Hewlett-Packard and IBM, that don't think it is working for them.

"Many of the rules of the game are being generated from successful areas such as the life sciences, but are applied across the campus

with little awareness of how unusual biotechnology is," says Woody Powell, a Stanford University sociologist who has followed the development of the biotech industry.

"University technology transfer is overly focused on the life sciences," agrees Lesa Mitchell, vice-president of the Ewing Marion Kauffman Foundation, a charitable trust based in Kansas City, Missouri, "and this has had a negative effect on university-industry collaborative models." The foundation is helping to support an effort by the Government-University-Industry Research Roundtable at the National Academies to try and find ways of improving the system.

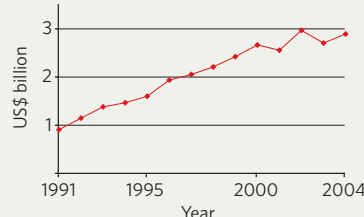
Levelling off

Stan Williams, head of quantum science research at Hewlett-Packard's laboratories in Palo Alto, California, says that universities have had trouble fulfilling their expectations for technology transfer into sectors such as semiconductors, in which it is relatively easy to invent around patent constraints.

"Going forward we could go from unfulfilled expectations to unmitigated disaster," he warns, given what he sees as the tendency of universities to overvalue their intellectual property. If a product requires dozens of patents, for example, and each university wants 5% of the profits, it soon becomes unfeasible to do the work, he adds.

"Fewer and fewer companies want to work with universities on sponsored research because they feel it doesn't make good business sense," says Susan Butts, external technology director at the Dow Chemical Company. "Companies could disadvantage themselves if it produces inventions that they are ultimately

INDUSTRIAL RESEARCH SPENDING AT US UNIVERSITIES



Locked in the lab: intellectual-property disputes make it hard for research to escape.

unable to license," she adds. Industry analysts point out that the growth in the flow of industry research dollars into universities has slowed and become more volatile in the past five years (see graph).

Complaints of troubled and lengthy negotiations prompted Butts and Bob Killoren, director of sponsored programmes at Pennsylvania State University, to approach the National Academies with the idea of hosting a project called Re-engineering Intellectual Property Rights Agreements in Industry-University Collaborations. With funding from Hewlett-Packard, IBM and foundations including the Kauffman, the project is intended to find ways of boosting collaborations between industry and universities in high technology.

After much vigorous debate, the project's participants feel they have found some common ground. They now intend to put together a model that can be readily customized to suit the requirements of different business sectors and project types. The resulting software package, dubbed 'TurboNegotiations', will be designed to help technology-transfer offices with the negotiation process — and keep a lid on legal bills. According to the Association of University Technology Managers (AUTM), universities spent \$205 million on patent

lawyers in 2003, of which \$86 million was recouped in legal fees from licensees.

Some industry watchers, such as Al Berkeley, former vice-chairman of NASDAQ, see the rifts between industry and academia as nothing new. The problem, as Berkeley told an AUTM meeting last year, is that half the technology-transfer job — discerning and marketing the potential revenue-generating patents — isn't being done. "They need to put just as much effort into marketing patents as investment banks put into marketing companies," he says.

Berkeley suggests that creating documents that point out key patent features, as well as where money could be made, would aid commercialization. To that end, the Kauffman Foundation released a web-based portal in September designed to help the research community connect with industry.

Drug problem

In the meantime, there are signs that technology transfer isn't even working well in the sector to whose needs it is most attuned — biotechnology. There are many possible reasons why fewer newly approved drugs have emerged from the system in the past ten years, including the fact that many of the more obvious approaches to finding drugs have already been explored.

Yet some critics of the system contend that the Bayh-Dole Act, which virtually gave birth to the US biotechnology industry, may now be strangling it, as universities seek patent protection on nearly everything (M. A. Heller & R. S. Eisenberg *Science* **280**, 698–701; 1998).

Leaders of medical schools and industry have acknowledged that intellectual-property issues are a roadblock to the drug-development process. At a January 2004 meeting hosted by the Food and Drug Administration, they agreed to take steps to better define, and find ways to share, 'precompetitive' research findings. They also promised to explore 'pooled' patents, already used by the information-technology industry, which grant broad access to entire suites of them at once.

"We've grown up expecting a certain return on investment in pharmaceuticals, and it isn't happening," says Frank Douglas, director of the newly established Center for Biomedical Innovation at the Massachusetts Institute of Technology. "People are beginning to question whether we're using the right model."

Douglas wants more precompetitive ideas and research tools to be shared freely, without the need for licensing. "If we could get people to agree on what is considered precompetitive," he says, "we might have a more rational approach to licensing fees and royalties." ■

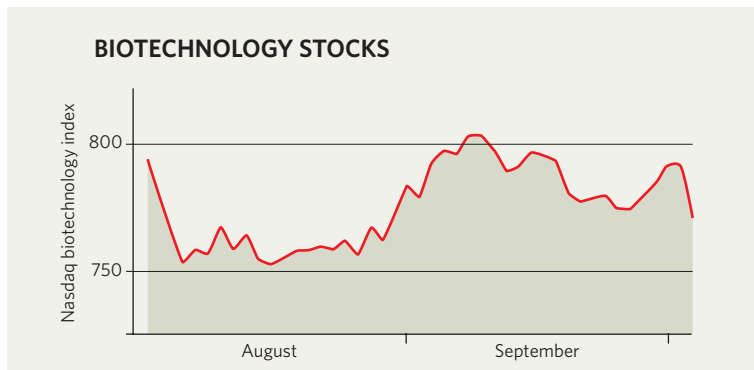
IN BRIEF

STEM-CELL SOURCE The National Institutes of Health has awarded \$16 million to the WiCell Research Institute at the University of Wisconsin to create and run the US government's first human national stem-cell bank. The grant will "dramatically reduce the cost of these cell lines to investigators", says James Thomson, WiCell's scientific director, who first isolated stem cells in 1998. Researchers have complained about the \$5,000 fee that WiCell currently charges to provide federally approved stem-cell lines (see *Nature* **435**, 272; 2005). This will now drop to \$500.

CHIPS IN Global semiconductor sales are growing rapidly and are set to comfortably exceed last year's record total. Sales this year up to the end of August totalled \$144 billion, the Semiconductor Industry Association says — 6% up on the same period in 2004. Semiconductor demand actually dropped in Europe, the United States and Japan, but grew sharply in the rest of Asia. Last year, the industry predicted that sales would flatten out in 2005.

NO CURE YET Human Genome Sciences of Rockville, Maryland, suffered a setback when it reported that LymphoStat-B, a drug it is developing to fight the autoimmune disease systemic lupus erythematosus, wasn't as effective as it had hoped during phase 2 trials. Shares in the 13-year-old company, which has yet to bring a drug to market, lost nearly a third of their value on the news. But the drug did subdue some symptoms of the disease, and Thomas Watkins, the company's chief executive, says he still hopes to move on to larger, phase 3 trials.

MARKET WATCH



This week Wood Mackenzie, an Edinburgh-based research and consulting firm, reviews recent trends in biotechnology stocks.

After a sustained increase from June, the Nasdaq biotechnology index was flat during the August holiday season. It then rose slowly in early September, driven by acquisition activity and better than expected earnings, before tailing off somewhat towards the end of the month.

Some US companies had a strong summer. MedImmune of Gaithersburg, Maryland, for example, regained the US marketing rights to Synagis, the company's respiratory-infection drug, from Abbott. The resulting increased earnings forecast ignited speculation that the company may be acquired, further boosting its share price. And California-based Gilead Sciences, meanwhile, continued to reap rewards from its flu treatment Tamiflu, which is marketed by Roche and is in high

demand because of fears of a global epidemic.

Chiron's share price rose after a \$4.5-billion acquisition offer by Novartis, even though the offer was rejected by Chiron's board (see *Nature* **437**, 317; 2005). And Eyetech Pharmaceuticals of New York saw its share price jump on the announcement of its acquisition by OSI Pharmaceuticals, also based in New York state. But OSI's share price fell sharply, because of fears among investors that the object of the deal, Eyetech's eye drug Macugen, will lose sales after 2006 following the anticipated approval of rival treatments.

Although the Nasdaq biotechnology index has recently outperformed more general indices, there is little evidence of any improvement in the financing environment. This suggests that mergers and acquisitions are likely to feature strongly in the industry's medium-term future.

► www.woodmac.com

Re-wilding: a bold plan that needs native megafauna

SIR — In their Commentary article “Re-wilding North America” (*Nature* **436**, 913–914; 2005), Josh Donlan and colleagues offer a vision of a North America populated with roaming throngs of megafauna such as cheetahs, elephants and tortoises. This is a welcome change from conservationists’ too often reactive and rearguard action against a tidal wave of human impacts and extinctions. But the proposal neglects people’s needs and political realities. Being bolder and more ambitious is the right idea — but Donlan and his colleagues have the wrong vision.

The Great Plains of North America are, in many areas, dominated by ranchlands and agricultural fields, as Steven Shay points out in Correspondence (*Nature* **437**, 476; 2005).

There are substantial social and economic problems associated with the presence of large animals on private lands, including, as the authors acknowledge, the enormous costs of building and maintaining fences to control them, as well as the environmental problems highlighted by Christopher Irwin Smith in Correspondence (*Nature* **437**, 318; 2005).

Most studies have shown that importing non-native species comes at huge economic and ecological cost.

A positive vision can catalyse a movement and generate real change, so it is critical for such visions to be grounded in reality.

As an alternative, we — 11 Smith post-doctoral fellows and three senior scientists with the Nature Conservancy (see https://webpace.utexas.edu/mas2687/SmithTNC_coauthors.html) — suggest that conservation efforts in North America should focus on restoring the megafauna native to this continent and ensuring that native species retain the evolutionary potential to adapt to novel environmental conditions, including those created by humans.

This approach, already adopted by the World Wildlife Fund and the American Prairie Foundation (see *Nature* **437**, 476; 2005), is more likely to be successful in the long run, as it respects the ecological conditions and environmental context to which these species are already adapted.

Martin A. Schlaepfer

(on behalf of TNC-Smith Fellows, classes 2003/4),
University of Texas, Austin, Texas 78712, USA

Evolution was fine, just not in the case of humans

SIR — Your News story “Fresh scope” (*Nature* **436**, 451; 2005) and Editorial “Keeping religion out of science class” (*Nature* **436**, 753; 2005) misrepresent the

Scopes case, or ‘monkey trial’. The substitute science teacher John Scopes was convicted and fined, not for teaching evolution in itself, but for his presentation of Darwin’s views on the descent of humanity.

In the 1920s, Tennessee, Arkansas and Mississippi passed laws banning the teaching of human evolution in public schools. According to Tennessee’s Butler Law, enacted in 1925, it was “unlawful ... to teach any theory that denies the story of the Divine Creation of man as taught in the Bible”. As noted by the Tennessee Supreme Court in the Scopes trial, “this enactment only intended to forbid teaching that men descended from a lower order of animals” (R. Moore, M. Jensen & J. Hatch *BioScience* **53**, 766–771; 2003).

After the Scopes trial, the laws banning the teaching of human evolution remained in effect for more than 40 years. But teaching students about Darwin’s general principle of evolution, with reference to non-human organisms, has never been illegal in the United States.

U. Kutschera

Institute of Biology, University of Kassel,
Heinrich-Plett-Strasse 40,
D-34109 Kassel, Germany

NIH moved quickly to help researchers after Katrina

SIR — I would like to add some details to your News story “New Orleans researchers fight to salvage work from submerged labs” (*Nature* **437**, 300; 2005) regarding efforts made by the National Institutes of Health (NIH) to assist its grantees in the Hurricane Katrina disaster zone.

Although it is true that the Federal Emergency Management Agency (FEMA) has primary responsibility for dealing with the immediate structural damage from a catastrophe — as I was quoted as saying — this quote does not do justice to the efforts by the NIH to promote a quick and full recovery from the storm. Leaders of damaged research institutions were contacted promptly and assured that the NIH would be at their sides in restoring facilities and research projects. The NIH immediately moved to inform grantees of the resources available, from grant extensions to temporary placements in the NIH intramural programme or other institutions.

The response to Katrina will require effort, flexibility and cooperation, but first and foremost we need profound commitment and solidarity. The NIH and the entire biomedical research community will stand together with those stricken by this disaster and by Hurricane Rita.

Elias Zerhouni

National Institutes of Health, 9000 Rockville Pike,
Bethesda, Maryland 20892, USA

Indian players in some of IT and biotech’s top teams

SIR — Your Outlook feature “Among the best: strong bonds” (*Nature* **436**, 494; 2005), quotes Govindarajan Padmanaban as saying “Indian scientists on the whole do not integrate in large groups”.

I disagree with this generalization. Indian scientists and engineers are widely employed across diversified disciplines, all over the heavily industrialized Western world. Many of today’s successful small and medium-sized IT and biotech companies were founded by Indian entrepreneurs, and elsewhere senior Indian executives have played key roles in running larger organizations.

A few familiar names in this ever-growing list are: Vinod Khosla, co-founder of Sun Microsystems; Sabeer Bhatia, founder of Hotmail; Arun Netravali, former president of Bell Labs; and Arun Sarin, chief executive of Vodafone.

The success of these companies, which to a large extent relies on teamwork and interpersonal skills, is, in itself, clear evidence that we Indians are great team players. Other contributory factors could be our familiarity with the English language (English is the *de facto* official language across India, thanks to the British Raj), our inborn spirituality and the extended-family culture that teaches cooperation, mutual respect and interdependence.

Mukund Mehrotra

Portola Pharmaceuticals,
270 East Grand Avenue,
South San Francisco, California 94080, USA

System to rank scientists was pedalled by Jeffreys

SIR — Jorge Hirsch’s index for measuring research achievement, as described in your News Story “Index aims for fair ranking of scientists” (*Nature* **436**, 900; 2005) was used by the geophysicist Harold Jeffreys for recording his cycling prowess, *n* being the highest number of days on which he had cycled *n* or more miles. I think he told me, some 35 years ago, that his *n* was 70 and that he first had the idea from his fellow cyclist, the astrophysicist Arthur Eddington.

A. W. F. Edwards

Gonville and Caius College,
University of Cambridge,
Cambridge CB2 1TA, UK

Contributions to Correspondence may be submitted to corres@nature.com. They should be no longer than 500 words, and ideally shorter. They should be signed by no more than three authors; preferably by one. Published contributions are edited.

BOOKS & ARTS

Before the storm

Katrina grabbed the headlines recently, but hurricanes have been a focus of attention for centuries.

Divine Wind: The History and Science of Hurricanes

by Kerry Emanuel

Oxford University Press: 2005. 304 pp.
\$45, £26.99

Howard B. Bluestein

The tropical atmosphere has recently been surpassing itself. In late August, Hurricane Katrina, following an encounter with South Florida, devastated parts of the US Gulf coast and led to the inundation of New Orleans. This hurricane has been called the worst natural disaster ever to hit the United States. The total number of fatalities has not yet been tallied, but it is expected to be high, and the number of people displaced, damage to property, and economic loss from the disappearance of businesses and interruption of oil and gas production are staggering. Hurricane Ophelia, which came soon after Katrina, brushed the Carolina coasts, and Hurricane Rita struck Louisiana and Texas. During the previous hurricane season, Florida was struck by a record number of hurricanes.

Divine Wind by Kerry Emanuel is therefore timely. Whereas other books on hurricanes focus on history, such as Erik Larson's *Isaac's Storm* (Little Brown, 1999), or on science, like Rick Anthes' *Tropical Cyclones* (American Meteorological Society, 1982), *Divine Wind* addresses both. Emanuel deftly interweaves an exposition of the science of hurricanes with historical accounts of major tropical cyclones and artists' impressions of the feelings that these tropical tempests instil. He also describes the challenge of predicting and understanding hurricanes. The book does a remarkable job of covering the history, science and terrible beauty of hurricanes, which makes it tempting not only to the storm aficionado and professional scientist, but also to the general public.

Emanuel's inclusion of art has a welcome humanizing effect. A scientific discussion of what happens when a hurricane makes landfall is framed within Richard Strauss's tone poem *Death and Transfiguration*. Technical figures such as satellite and radar images are intermixed with historical photographs and realistic, abstract and sometimes apocalyptic paintings. There are poems of the regular, not tone, variety, and colourful and easy-to-interpret figures illustrating scientific principles. Instrumentation such as radar, satellites,

aircraft, remotely piloted aircraft and scatterometers is also described and explained. Throughout, Emanuel achieves a perfect balance between the scientific and non-scientific aspects of hurricanes.

My favourite chapter is a photo essay depicting the inside of a hurricane or typhoon. Hurricanes are dangerous, but they are beautiful too, if observed from the safety of an aircraft inside the eye of the storm or from space. One photograph, taken during a US Air Force reconnaissance flight, shows the eye of a typhoon viewed from below. Others have views looking straight down at the turbulent ocean surface and windblown spray, and there is even one image from the space shuttle. Emanuel compares a view of the eye of a typhoon with a scene from Dante's *Inferno*.

There are accounts of major historical hurricanes and typhoons. Some of them affected the course of history, such as the typhoons that struck Japan in the thirteenth century (the 'divine wind', or *kamikaze*, that saved Japan from Kublai Khan), and a hurricane in the sixteenth century that interfered with an attempt by the French to settle Florida. The 'great hurricane' of 1780 in the Caribbean was one of the deadliest hurricanes ever to strike in that part of the world. The book also covers the 1900 Galveston hurricane, the New England

hurricane of 1938, Hurricane Camille in 1969, Cyclone Tracy in Australia in Christmas 1974 and Hurricane Andrew in 1992. The effects of Katrina are likely to be as long lasting, politically, socially and economically, as those of its predecessors. Other books have dealt with the historical accounts in more depth, but the overall effect of those in *Divine Wind* is unique.

Emanuel notes a similarity between *The Tempest* and an account of a hurricane in 1609 given by William Strachey, an acquaintance of William Shakespeare. Many scholars say this rules out Edward De Vere as a possible author of *The Tempest*, because he died before Strachey told his tale. Strachey's account itself is poetic:

For four and twenty hours the storm in a restless tumult, had blown so exceedingly, as we could not apprehend in our imagination any possibility of greater violence, yet did we still find it, not more terrible, but more constant, fury added to fury, and one storm urging a second more outrageous than the former... The Sea swelled above the clouds, and gave battle unto heaven.

A pioneering hurricane researcher, Emanuel presents the science at a level that is not too technical for non-specialists, yet is sufficient

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Tossed and blown: ships take a battering in Engel Hoogerheyden's 1795 painting *The Eye of the Hurricane*.

to describe the basic physics with few equations. Hurricane formation, energetics and ocean interaction are all clearly explained. The fundamentals of numerical weather prediction using computers, including chaos theory and its relevance to forecasting, are also well treated. The only minor criticism I have is that each of the forces that control the behaviour of hurricanes is treated separately. It would have been nice if the effects of all the forces had been summarized, particularly in the section on numerical weather prediction,

where only the sum of all the forces is noted.

Emanuel has recently been in the public eye as a result of his recent letter in *Nature* (436, 686–688, 2005), in which he suggested that any further warming of the troposphere might increase the destructive potential of tropical cyclones. His book ends with some further provocative thoughts on hurricanes and climate. ■

Howard B. Bluestein is at the School of Meteorology, University of Oklahoma, Norman, Oklahoma 73019, USA.

Fetal affliction

The Fetal Matrix: Evolution, Development and Disease

by Peter Gluckman & Mark Hanson

Cambridge University Press: 2005. 272 pp. £50, \$90 (hbk); £24.99, \$39.99 (pbk)

Michael Sargent

The idea of the womb as a tranquil nursery, untroubled by worldly pressures, is charming but wrong: the unborn baby reacts and adapts to its unique environment, with profound consequences for later life. This is the view that emerged in the 1990s from some remarkable investigations led by David Barker at the University of Southampton. Retrospective studies of cohorts of people born in the 1920s indicated that babies that were small for their gestational age were more susceptible than average in middle age to coronary heart disease, type 2 diabetes, hypertension and osteoporosis.

The Fetal Matrix by Peter Gluckman and Mark Hanson is a fascinating and important book about the responses of the mammalian fetus to its environment. Building on Barker's proposals, the authors suggest that the fetus can respond to a potentially harsh nutritional landscape by scaling down the developmental enterprise to create a 'survival phenotype' — a small, lean body with undiminished capacity for reproduction. Visceral organs are underdeveloped relative to the brain because fewer cells are allocated to capillaries in muscle, to nephrons, to the liver or to insulin secretion. Energy is conserved by avoiding unnecessary muscle growth and using any excess to create fat deposits that can be mobilized when food is scarce. For most of human history, the survival phenotype has been an important and appropriate adaptation when food and population were in a delicate malthusian balance.

About sixty years ago, when food became available in unprecedented abundance in the developed world and work became less physically demanding, the survival phenotype began to be a liability for some people. Those affected were likely to develop abdominal fat, high blood pressure and reduced responsiveness to a glucose surge (insulin resistance)

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Womb for improvement? Events before birth can have huge consequences for health in later life.

— all signs of the 'metabolic syndrome'.

Barker's view of the fetal origins of disease is not universally accepted. Unfortunately, readers of this book who want to disentangle the arguments for themselves will be frustrated by the unsatisfactory link between the text and specific references. But compelling support for the concept, less dependent on fine statistical judgements, comes from work using experimental animals. This shows that the survival phenotype and its pathological consequences, hypertension and insulin resistance, can be induced *in utero* when the mother's diet is deficient in protein or micronutrients, or by treatment with the stress hormone cortisol.

The survival phenotype can be provoked in the fetus by undernutrition, by the physical constraint imposed by having a small mother, or by cortisol crossing the placenta causing growth restriction and accelerated maturation. A constraint on birth weight is inevitable in small mothers and is repeated again when their daughters (also small) become pregnant. This epigenetic phenomenon is seen with special poignancy in some parts of India, where persistent malnutrition generates a population of phenotypically small mothers whose babies

are among the smallest known. The long-term effect on unborn babies of a period of malnutrition of precise intensity and duration is best known from the Dutch *Hungerwinter* of 1944–45. Many babies conceived during this episode developed the survival phenotype but were often only marginally underweight.

What provokes the survival phenotype? Maternal nutritional deficiency or exposure to cortisol causes the promoters of certain genes to remain unmethylated in the early embryo, precipitating changes of gene expression that confer an altered phenotype that persists into adult life. These genes affect activities as diverse as apoptotic remodelling of the embryo and the capacity to groom offspring, and it is a fair bet that more important examples will emerge. Offspring with the survival phenotype that gain weight quickly during childhood are prone, as adults, to the characteristic pathology of the metabolic syndrome. In countries that are now undergoing rapid economic development, a serious issue of public health is unfolding that is likely to affect people whose standard of living is increasing. Indeed, the World Health Organization has predicted that in 20 years' time, 5% of the world's population will develop type 2 diabetes.

This book is a thought-provoking account of a topic in which developmental biology, physiology and clinical medicine intersect. It focuses on the idea that the development of the mammalian fetus normally anticipates the postnatal environment, triggering a "predictive adaptive response", but may misjudge the situation. A fetus with the survival phenotype that enters an unexpectedly bounteous world starts on a developmental trajectory that may predispose the individual to ill health.

The authors convey admirably the physiological implications of an important idea, but they are less forthcoming about other dimensions of the subject that might appeal to a non-specialist reader. For example, the literature suggests that the status of the immune system and the inflammatory response, and possibly certain behavioural and psychological characteristics, are affected by events in the uterus. Similarly, investigations by historians into trends in human physique are not included. These are immensely interesting records that are an important barometer of life experience resulting from physiological decisions taken *in utero*.

I was glad to learn that 'fetus' — Latin for 'offspring', and used in English since the fourteenth century — is the correct usage and that 'foetus' is a recent pseudo-Greek affectation. More obscurely, the 'matrix' of the title alludes to a surprising dictionary definition (it means 'womb') and perhaps to the set of variables that can have such far-reaching consequences for the fetus. ■

Michael Sargent is a developmental biologist at the National Institute for Medical Research, Mill Hill, London and the author of *Biomedicine and the Human Condition: Challenges, Risks and Rewards*.

Nuclear reactions

Wormwood Forest: A Natural History of Chernobyl

by Mary Mycio

Joseph Henry Press: 2005. 286 pp. \$27.95

Brenda Howard

The Chernobyl nuclear accident in April 1986 was swiftly followed by a large-scale evacuation of an area around the plant, including territory in both Ukraine and Belarus. Although the rural residents and inhabitants of the major town of Pripjat were evacuated, the exclusion zone is still occupied by many thousands of people. Most of them are associated with the continuing activities at the Chernobyl nuclear facilities, but some are rural residents who have returned to live in their previously abandoned villages.

In *Wormwood Forest*, Mary Mycio, a journalist with an ethnic Ukrainian background, provides the reader with a vivid impression of what the exclusion zone around the Chernobyl plant is really like. She first visited the zone ten years after the accident, and describes the many people who assisted her on her visits, the local people she met, and the various bureaucratic niceties involved in administering and visiting the zone.

The book starts by correcting a commonly held mistaken impression that Chernobyl takes its name from the Ukrainian word for wormwood, a medicinal herb. *Chornobyl* is actually mugwort, *Artemisia vulgaris*, not wormwood, which is *A. absinthium*. It is odd, then, that the book is titled *Wormwood Forest*

and has an associated quote on the cover: "And the name of the star is called wormwood; and the third part of the waters became wormwood; and many men died of the waters because they were made bitter."

A large part of the book focuses on the wildlife that has flourished in the exclusion zone since the removal of human influence, and describes how the extent of radionuclide contamination varies with species. Mycio describes her encounters with many animals including storks, deer, moose, wild boar and the introduced Przewalski's horses. Although the focus is on the natural history of the zone, other key issues addressed include radioactive-waste problems, concerns about the various water bodies, and the radiological consequences of the accident.

The author has clearly made a considerable effort to understand the complex social and scientific issues connected with the region and has managed to explain them to the lay reader in a refreshingly clear, yet interesting, style. The blend of social comment, personal impressions and science is unusual and makes for an informative read. Her explanations of the basics of radioactivity and the issues relating to the accident and its consequences are mostly sound and easy to understand.

For scientists with some knowledge of radioecology, however, her attempt to describe some of the issues surrounding radioactive contamination in plain (and sometimes rather colourful!) language has led to a few conclusions being too generalized. Many radioecolo-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

All change: contamination from Chernobyl has led to defects in a range of plants and animals.

gists might disagree with some of her rather sweeping statements about the importance of processes such as resuspension, fire or the lateral transport of radioactivity. Some statements by people quoted in the book seem to be incorrect, although they are at least presented as quotes. And many scientists will be frustrated by the decision not to use references when discussing scientific studies, as there are always some studies of which you are unaware (especially in the Russian-language literature).

There is great variability in the extent to which different ecosystem components have been contaminated by the dominant long-lived radionuclides radiocaesium, radiostrontium and plutonium, as the author acknowledges. However, she sometimes refers to single measurements with no indication of the associated error (probably because in many cases it doesn't exist). Conclusions based on such limited information are inevitably susceptible to criticism.

For much of the book, Mycio refers to readings of the external dose and to the colours of contamination maps to give an impression of the extent of contamination of each area. She often comments on how 'high' or 'low' the readings are and whether she feels this to be safe or not. She uses a similar approach when considering radioactivity in a food product, often referring to the national limits on the number of becquerels used by the Ukrainian and Belarusian authorities. These limits are lower than internationally agreed limits and those of many other countries, so it would have been useful to have more information about them to provide context for readers.

This book is not a key reference source for information on radionuclide contamination of the environment close to the Chernobyl reactor, but then it doesn't claim to be. It is very much a personal reflection that successfully debunks many of the more outrageous myths and rumours about the region, and is an interesting and mostly enjoyable read. ■

Brenda Howard is at the Centre for Ecology and Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK.



EXHIBITION

Collectors' items

In the exhibition *Saved by Science*, Justine Cooper takes us on a voyeuristic journey through the labyrinth of vaults in the American Museum of Natural History in New York. The photographs and video footage capture a behind-the-scenes glimpse of one of the world's largest and most valuable scientific collections. Many of the specimens seen are not normally on public display.

Some of the images are beautiful, others grotesque — such as the monkey fetuses curled in mock sleeping position in preservation jars.

The museum, with its 25 interconnected buildings and vast vaults, spans four blocks. A five-minute video wanders through corridors dotted with cabinets of specimens, and a soundtrack of screeches, chirps and babblings seems to bring the specimens — such as the yellow honeyeaters shown here — to life.

Saved by Science can be seen at the Mary Place Gallery in Sydney, Australia, from 25 October to 6 November 2005. **C.D.**

NEWS & VIEWS



PALAEOANTHROPOLOGY

Further fossil finds from Flores

Daniel E. Lieberman

New fossil discoveries on Flores, Indonesia, bolster the evidence that *Homo floresiensis* was a dwarfed human species that lived at the end of the last ice age. But the species' evolutionary origins remain obscure.

When Gulliver was shipwrecked on the East Indian island of Lilliput in Swift's satirical novel, he was astonished to discover tiny humans. Last year's announcement^{1,2} of a newly discovered species of tiny human from the Indonesian island of Flores was far more astonishing, because it wasn't made up. This intriguing scientific story continues on page 1012 of this issue³, where Morwood and colleagues describe further fossil evidence from the cave of Liang Bua on Flores.

The original fossil remains¹ consisted primarily of a single partial skeleton (LB1), excavated from deposits in Liang Bua dated to the end of the last ice age. Stone tools, evidence of fire-making and the bones of a dwarfed elephant species were also found, those bones apparently being the result of hunting. The LB1 skeleton, dated to 18,000 years ago, was probably a female, just over a metre tall. It had a brain volume of 380 cm³, roughly the size of a chimpanzee brain. Although LB1 has a somewhat primitively shaped pelvis, it shares many derived characteristics of the genus *Homo*, particularly in the teeth, jaw and cranium. These similarities, combined with other distinctive features, led Brown and colleagues¹ to propose a new species, *Homo floresiensis*. They further suggested that *H. floresiensis* was a dwarfed descendant of *Homo erectus*, another hominid species, which is thought to have arrived on Flores by 800,000 years ago⁴.

Homo floresiensis caused a stir by challenging preconceptions. If it is a new species, then we shared this planet with other hominids

much more recently than anyone thought — long after the Neanderthals became extinct, after modern humans arrived in Australia, and at about the time that agriculture was first invented. More unusual is the proposal that *H. floresiensis* evolved from *H. erectus* through dwarfing. This phenomenon, known as endemic or island dwarfing, sometimes occurs on islands when species are released from the pressures of predation but become constrained by limited resources and small population sizes⁵. In such conditions, large animals tend to become smaller and small animals tend to become larger. The process was clearly occurring on Flores, whose fauna includes giant rats and now-extinct miniature elephants. What captures the imagination is that dwarfing might have occurred in humans, who often buffer themselves from natural selection through cultural means such as tool production and fire-making, both evident at Liang Bua².

The Liang Bua finds have generated controversy. Two alternative hypotheses, yet to be published in the peer-reviewed literature, have been proposed. One is that the LB1 skeleton is a pygmy human, not a new hominid species. The other is that LB1 is a human who suffered from a form of microcephaly, a pathological condition characterized by an abnormally small brain and head, and which can also cause dwarfism^{6,7}.

Morwood and colleagues³ now counter some of these claims with evidence recovered during excavations in 2004. The material

substantially expands the sample attributed to *H. floresiensis*, and provides additional details about the proposed species. The new fossils consist of the right humerus, radius and ulna of the LB1 skeleton, the mandible of a second individual (LB6), and assorted other remains including two tibiae, a femur, two radii, an ulna, a scapula, a vertebra, and various toe and finger bones. The researchers think that the sample includes the remains of at least nine individuals.

The analysis³ focuses on the new mandible (LB6), a new tibia (LB8) and the LB1 skeleton's reunited arm bones. Of the many details, several merit special attention. First, the new mandible is extraordinarily similar to the first one. They almost certainly belong to the same species. Both mandibles share distinctive dental features, and they lack chins — a chin being a unique feature of all *Homo sapiens* regardless of their stature, including most microcephalics (Fig. 1, overleaf). In addition, the new tibia and arm-bone fossils not only confirm that the Liang Bua hominids were short, about a metre tall, but also indicate that they had relatively long arms. In many ways, the LB1 skeleton's body proportions are less like any adult human's, including adult pygmies, than those of an australopithecine — an earlier hominid lineage, thought to have been confined to Africa.

Another notable point is that the Liang Bua fossils come from a lengthy temporal span during which the cave's inhabitants were hunting animals, producing stone tools and

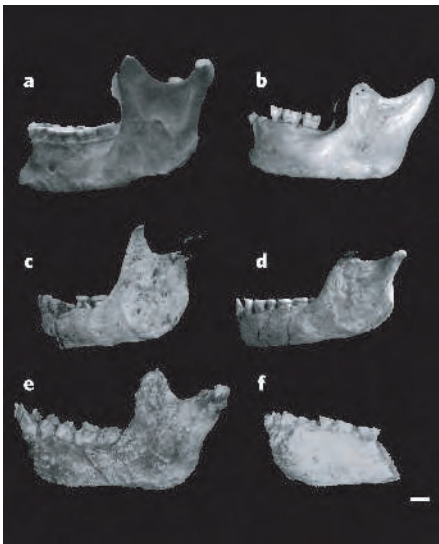


Figure 1 | Comparison of mandibles. **a**, *Homo sapiens* from Upper Cave Zhoukoudian, China; **b**, a microcephalic *H. sapiens* from Mauritius (Peabody Museum, Harvard Univ.); **c**, *Homo floresiensis* (LB1); **d**, *H. floresiensis* (LB6); **e**, early African *Homo erectus* (KNM-WT 15000); **f**, *Australopithecus afarensis* (Laetoli Hominid 4). The Liang Bua mandibles lack a chin, unlike those of *H. sapiens*, including the microcephalic example shown here. Lack of a chin is an ancestral feature of hominids, and is also a characteristic of *H. erectus* and *A. afarensis*. Scale bar, 1 cm.

making fire. Although the original LB1 skeleton is estimated to be 18,000 years old, a child's radius was found in deposits estimated to be 12,000 years old, and the new mandible is estimated to be 15,000 years old; other finds may be as old as 95,000 years^{2,3}.

The fossils also all seem to be similarly small, refuting the contention that the LB1 skeleton was simply an aberrantly dwarfed, pathological specimen. If they were pathological, then the Liang Bua fossils would have had to have come from a population of short, microcephalic humans that survived for a long time, or one that was susceptible to high frequencies of microcephaly and dwarfism. Such possibilities strain credulity; moreover, a three-dimensional analysis of the LB1 brain-cast⁸ found the brain to be unlike a microcephalic's, and more like that of *H. erectus* than *H. sapiens*. Microcephaly, however, can have many causes⁹, and further studies that use larger sample sizes and analyse a wide range of syndromes will be necessary to test the hypothesis completely.

All in all, it seems reasonable for Morwood and colleagues³ to stick to their original hypothesis that *H. floresiensis* is a new species. But they are less certain about whether it evolved from *H. erectus* or from some other species, and raise the possibility that the species derives from an unknown small-bodied hominid, more primitive than *H. erectus* and with australopithecine-like body proportions. This seems unlikely, given the many derived features characteristic of *Homo* present in

H. floresiensis. However, the variability evident in new fossil material of early *Homo* from Georgia¹⁰ and Kenya¹¹ underscores just how little we know about diversity within the human genus.

What is needed to test the various proposed hypotheses and convince the sceptics? As always, more fossils would help. Because Flores was inhabited at least 800,000 years ago⁴, it will be useful to find older fossils and see if they look like *H. erectus*, or something else. Further, if the island-dwarfing hypothesis is correct, then the island's earliest inhabitants should be larger than the Liang Bua fossils; and if dwarfing occurred gradually, then it might even be possible to find fossils intermediate in size and shape between *H. floresiensis* and its ancestor. More evidence on when *H. sapiens* first arrived on Flores is also needed.

Such fossils may not be there to be found, however, and testing evolutionary hypotheses from even a well-sampled fossil record poses many challenges. One obvious avenue is to apply morphometric methods that analyse three-dimensional shape independently of size, to test whether the fossils are scaled down versions of *H. erectus*, *H. sapiens* or something else¹². Many syndromes can cause microcephaly and dwarfism⁹, and they all need to be considered. Finally, it would be interesting to find out how the hypothesized size reductions

on Flores differ from those of other dwarfed island mammals.

The finds from Liang Bua are not only astonishing, but also exciting because of the questions they raise. More analyses that may lead to answers are on the way, despite an unfortunate custody dispute over the fossils that led to some being damaged¹³. Eventually, we might even reach a consensus about the nature and origins of *H. floresiensis*, particularly after the data become available to the general scientific community.

Daniel E. Lieberman is at the Peabody Museum, Harvard University, 11 Divinity Avenue, Cambridge, Massachusetts 02138, USA. e-mail: danlieb@fas.harvard.edu

1. Brown, P. *et al.* *Nature* **431**, 1055–1061 (2004).
2. Morwood, M. J. *et al.* *Nature* **431**, 1087–1091 (2004).
3. Morwood, M. J. *et al.* *Nature* **437**, 1012–1017 (2005).
4. Morwood, M. J., O'Sullivan, P. B., Aziz, F. & Raza, A. *Nature* **392**, 173–176 (1998).
5. Foster, J. B. *Nature* **202**, 234–235 (1964).
6. Henneberg, M. & Thorne, A. *Before Farming* **4**, 2–3 (2004).
7. Dalton, R. *Nature* **434**, 432–434 (2005).
8. Falk, D. *et al.* *Science* **308**, 242–245 (2005).
9. Gilbert, S. L., Dobyns, W. B. & Lahn, B. T. *Nature Rev. Genet.* **6**, 581–590 (2005).
10. Vekua, A. *et al.* *Science* **297**, 85–89 (2002).
11. Spoor, F., Leakey, L. N. & Leakey, M. G. *Am. J. Phys. Anthropol. Suppl.* **40**, 195–196 (2005).
12. Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. *Geometric Morphometrics for Biologists: A Primer* (Elsevier, San Diego, 2004).
13. Culotta, E. *Science* **307**, 1848 (2005).

PLANETARY SCIENCE

The impact of Deep Impact

Paul D. Feldman

A good look at the Deep Impact cometary encounter was taken by the Rosetta mission, itself on the way to a rendezvous with a comet in 2014. So what is a comet — icy dustball or dusty iceball?

NASA's Deep Impact mission aimed to bring a spacecraft weighing 362 kilograms into collision with the periodic comet 9P/Tempel 1 on 4 July 2005. The mission was, by all accounts, a smashing success: the images returned by both the impactor, which showed the first detailed views of a cometary nucleus from just before collision, and the mother-ship, showing the impact and its immediate after-effects (Fig. 1), are truly spectacular¹.

The scientific success of Deep Impact depended not just on the accuracy of the spacecraft's controllers, but also on an extensive network of Earth- and space-based observers that appropriated almost all of the planet's astronomical resources to study the effects of the collision in every region of the electromagnetic spectrum². The first results from this wealth of data have already been presented at conferences in Búzios, Brazil³, and Cambridge, UK⁴. On page 987 of this issue, Küppers *et al.*⁵ present observations of Deep

Impact from the European Space Agency's Rosetta spacecraft, launched on 2 March 2004. This spacecraft is itself on a ten-year voyage to another periodic comet, 67P/Churyumov-Gerasimenko, where it will orbit the comet's nucleus and, in November 2014, deposit a lander to probe the composition and physical nature of its surface and immediate subsurface.

Comets are of interest because their composition is expected to reflect conditions that were prevalent when the Solar System was formed. So far, only the make-up of a comet's enveloping coma is known to any great extent. Within about 3 AU of the Sun (1 AU, or astronomical unit, is the distance from Earth to the Sun), the gaseous component of this unbound atmosphere is dominated by water that has been sublimated (converted directly from ice to vapour) from the surface of the comet's nucleus by the warming effect of the Sun. Dust grains are also present in the coma and tail,

and the reflection of solar radiation off them produces most of a comet's visible light. The ratio of ice to dust in a comet's interior is assumed to be quite different from its ambient-surface value: hence the expected value of Deep Impact⁶, which allowed the investigation of frozen residues of primordial ice liberated from the interior of Tempel 1.

The Rosetta probe observed the Deep Impact encounter using OSIRIS — its Optical, Spectroscopic, and InfraRed Imaging System⁵. This suite of two cameras has been designed for in-flight mapping of Churyumov–Gerasimenko's nucleus and inner coma of gas and dust — among other things, to help identify good landing sites for the probe. Although lacking the sensitivity achievable with the large collecting area of an astronomical telescope, the OSIRIS cameras possessed three clear advantages over Earth-based instruments for observing the Deep Impact collision. First, they were almost twice as close to Tempel 1 as was Earth at the time of impact. Second, they afforded a continuous view from six days before to ten days after the event. Third, the cameras' extraterrestrial position allowed them to measure, through a filtered channel, the spectral line of the hydroxyl (OH) radical at a wavelength of 308 nanometres, without the need to correct for attenuation in Earth's atmosphere.

This hydroxyl measurement is particularly pertinent to establishing the amount of water ice contained in the Tempel 1 residues. Light from the Sun breaks down about 85% of the water in a comet's coma into OH and atomic hydrogen; most of the OH is then itself photodissociated into atomic oxygen and hydrogen. The rates at which these processes occur in the solar radiation field are well known⁷: at 1.506 AU, the distance of Tempel 1 from the Sun at the time of impact, the average lifetime of a water molecule until photodissociation is about two days, and that of a OH radical slightly longer. (OH, normally highly reactive, is becalmed in space in the absence of material to react with.) This meant that Küppers *et al.*⁵ could use their relatively wide-field images to watch the build-up of OH and to determine, independently of kinematic models of the outflowing gas, the total amount of water excavated by the impact.

The total water abundance could also have been detected directly at infrared frequencies by Deep Impact, but the viewing geometry of the craft's infrared spectrometer made a quantitative assessment difficult¹. Terrestrial spectroscopic observations were less suited to the task of assessing the water outflow, as the necessary high spectral resolution requires a narrow viewing slit that would have excluded a large fraction of the outwardly expanding emitting region⁸.

Küppers *et al.*⁵ found the total amount of water released from Tempel 1 to be just under



Figure 1 | Big Bang. The aftermath of the collision of the Deep Impact probe with comet 9P/Tempel 1, seen 67 seconds later by the High Resolution Imager aboard the mother spacecraft. (The Tempel 1 comet measures roughly $4 \times 4 \times 14$ kilometres.)

5,000 tonnes. Determining the mass of dust liberated — the second part of the puzzle — with the Rosetta images is, however, less straightforward. By assuming certain optical parameters, such as the proportion of incident radiation that scatters off the dust rather than being absorbed (the 'albedo'), the total area of the material added to the coma by the impact can be derived. The conversion to volume (and so mass) depends sensitively on the density of the dust grains and their size distribution. Taking a wide range of possible values for these parameters, Küppers and colleagues find a dust–ice ratio of greater than one. But their spread of values for the total dust mass is significantly higher than results from two sets of ground-based spectroscopic measurements in the mid-infrared^{9,10}. The

values estimated from these measurements are below 1,000 tonnes, and so imply a much lower dust–ice ratio. The infrared results also depend on models of grain size and density, but are more reliable as they are derived from thermal-emission measurements of larger grains that contain a greater fraction of the total dust mass. Of course, the total mass of material liberated by the impact could have been estimated quite simply from the dimensions of the crater left behind — had not the cloud of fine dust that Deep Impact kicked up obscured the probe's own view.

Issues of the dust and water composition of Tempel 1 will undoubtedly provoke lively debate over the next few months as more analyses, including those from Deep Impact, are completed and made public. At this point, it seems that we have only scratched the surface of this comet, and that many of the conclusions about its make-up are likely to change. Some questions may even remain unanswered until Rosetta reaches its ultimate target.

Paul D. Feldman is in the Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA.
e-mail: pdf@pha.jhu.edu

1. A'Hearn, M. F. *et al.* *Science* advance online publication doi:10.1126/science.1118923 (2005).
2. Meech, K. J. *et al.* *Science* advance online publication doi:10.1126/science.1118978 (2005).
3. www.on.br/acm2005
4. www-outreach.phy.cam.ac.uk/dps2005/index.html
5. Küppers, M. *et al.* *Nature* **437**, 987–990 (2005).
6. A'Hearn, M. F. *et al.* *Space Sci. Rev.* **117**, 1–21 (2005).
7. Budzien, S. A., Festou, M. C. & Feldman, P. D. *Icarus* **107**, 164–188 (1994).
8. Mumma, M. J. *et al.* *Science* advance online publication doi:10.1126/science.1119337 (2005).
9. Harker, D. E. *et al.* *Science* advance online publication doi:10.1126/science.1119143 (2005).
10. Sugita, S. *et al.* *Science* advance online publication doi:10.1126/science.1119091 (2005).

ECOLOGY

Roots of stability

Peter D. Moore

The 'insurance hypothesis' holds that ecosystem diversity is a good thing because diversity confers overall stability in the face of stressful conditions. Experiments on grassland support that view.

Most ecologists agree that diversity is a desirable attribute for an ecosystem. But not all of them agree on why that is. It would be satisfying to be able to claim that diversity makes an ecosystem more stable, implying that a diverse system is better able to resist, or recover from, disturbance. But testing the point experimentally is no simple task.

As they report in *Functional Ecology*, Kahmen and colleagues¹ have taken a step in this direction. They show that when grassland

with a high plant diversity is experimentally stressed by drought, it develops rooting systems more effectively than grassland with low diversity.

The debate about diversity and stability has swung back and forth violently, depending on whether arguments are based on intuitive reasoning, theoretical modelling or experiments. The intuitive approach suggests that a complex ecosystem offers more options than a simpler one when placed under stress. Prey-switching

in food webs could reduce the likelihood of over-exploitation of a particular resource leading to species extinction; alternatively, a multitude of physiological adaptations among the members of a community would give the whole ecosystem a better chance of coping with catastrophe.

This type of reasoning is known as the 'insurance hypothesis'. It went out of favour for many years, because theoretical modelling² seemed to contradict the gut feelings of ecologists. Many models indicated that high levels of complexity could lead to dynamic fragility, meaning that the ecosystem would be stable only within a very limited range of environmental conditions. Controlled field observations³, coupled with laboratory and field experiments⁴, have provided a possible means of settling the question. But simple conclusions have been elusive, partly because measuring stability, or even defining it, has proved difficult.

One way forward is to look at the influence of diversity on how efficiently an ecosystem functions⁵. For example, it is relatively easy to measure plant productivity, especially above the soil surface. This is a fundamental process in an ecosystem on which the flow of energy to other ecosystem components ultimately depends. So it is reasonable to think that the stability of the ecosystem depends on the maintenance of this function. The results from studies involving ecosystem diversity and productivity have differed, however; sometimes diversity seems to have led to increased productivity⁶ and sometimes to inherent instability⁷.

Kahmen and colleagues¹ have also adopted this functional approach in their studies of grassland stability in central Germany. They selected 19 sites within mountain grassland, all situated on the same soil type but differing in their plant diversity. The sites had a similar land-use history, having been ungrazed but harvested for hay twice a year for the past ten years. Within each site, control plots were matched with plots in which a transparent cover prevented rainfall reaching the ground from April to June, thus simulating drought. The use of drought as a stress factor is ecologically reasonable given the recent climatic trends towards drier summers and, especially, the increasing frequency of short-term episodes of extreme drought.

Kahmen *et al.* subsequently measured above-ground productivity, and also used soil-coring techniques to measure the productivity of the root systems. They found no relationship between grassland diversity (estimated by a measure known as the Shannon index) and above-ground productivity. But they show that below-ground productivity increased as a response to drought in the more diverse sites. Under these stressed conditions, the plant community as a whole is evidently diverting more of its fixed carbon to root production than to shoot production.

In Kahmen and colleagues' experiment, it

was not possible to separate root biomass into its component species, so there is no information about which species of plants are shifting their resources to root growth. But it seems likely that in a diverse system there is a greater probability that some species react to drought by changing their resource-allocation strategies. Such a change is clearly advantageous to individual plants because it gives them a better chance of competing for the limited water in the stressed conditions. Viewing the community as a whole, therefore, one can argue that a diverse system will be better able to cope with drought and to survive the experience relatively intact.

The insurance hypothesis thus gains support from this study. Beyond that, plant ecologists are again reminded that processes

occurring underground, although less apparent, matter as much as those occurring above the soil surface. ■

Peter D. Moore is in the Department of Biochemistry, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, UK.
e-mail: peter.moore@kcl.ac.uk

1. Kahmen, A., Perner, J. & Buchmann, N. *Funct. Ecol.* **19**, 291–298 (2005).
2. May, R. M. in *Theoretical Ecology* (ed. May, R. M.) 142–162 (Blackwell Scientific, Oxford, 1976).
3. Frank, D. A. & McNaughton, S. J. *Oikos* **62**, 360–362 (1991).
4. Tilman, D. & Downing, J. A. *Nature* **367**, 363–365 (1994).
5. Loreau, M., Naeem, S. & Inchausti, P. (eds) *Biodiversity and Ecosystem Functioning* (Oxford Univ. Press, 2002).
6. Tilman, D., Wedin, D. & Knops, J. *Nature* **379**, 718–720 (1996).
7. Pfisterer, A. B. & Schmid, B. *Nature* **416**, 84–86 (2002).

MATERIALS SCIENCE

At a stretch

A polymer based on the elastic protein that enables fleas to perform their extraordinary jumping feats has been synthesized. The material, described by C. M. Elvin *et al.* in this issue (*Nature* **437**, 999–1002; 2005), is, perhaps unsurprisingly, rubbery and highly resilient; indeed, some of its properties exceed those of a material used to make bouncy balls for the playground.

The elastic protein resilin, the source of the authors' inspiration, comes from the same family of biomacromolecules as spider silk and elastin (a protein that makes animal tissues flexible). It is, in fact, found in most insects, where it facilitates not just jumping, but also many actions requiring efficient energy storage and rapid repetitive movement — the chirping of cicadas and the flapping of dragonflies' wings being other examples.

Elvin *et al.* produced their polymer by first cloning part of a *Drosophila* fruitfly gene and expressing it in *Escherichia coli* bacteria. They thus produced large quantities of a peptide found in a precursor molecule to resilin that contains 17 repeats of an amino-acid sequence thought to be responsible for elasticity. A photochemical reaction established crosslinks in the



peptide to produce a solid material.

The result, a thread about one millimetre across, glowed blue when bathed in ultraviolet light (see image), and was easily cast into a range of shapes. Its ability to recover from deformation (the property known as resilience) was as good as that of natural resilin taken from a wing tendon of a dragonfly. It was also similar to that reported for elastin, and far higher than that of several synthetic rubbery polymers. A strip of the crosslinked resilin could be stretched to more than three

times its original length without breaking.

The authors attribute the mechanical properties of the polymer to the three-dimensional amorphous nature of the crosslinked protein matrix. They also believe that a clearer understanding of the part played by water as a solvent and plasticizer could lead to the development of a whole new range of rubbery materials. Meanwhile, they suggest, their jumping-flea polymer could be a leap forward for biomedical and other engineering fields.

Rosamund Daw

DEVICE PHYSICS

No-nuisance noise

Adi R. Bulsara

'Silence is golden' is a maxim of limited applicability where stochastic resonance holds sway. The effect uses noise to boost signal output in certain systems — and has just been seen in oscillators on a very small scale.

Stochastic resonance^{1,2} encapsulates the sexy notion that moderate (and, ideally, carefully controlled) levels of noise in a nonlinear dynamical system can actually enhance the information throughput — and so improve the sensing and processing of otherwise undetectable signals. Originally postulated as a mechanism to explain how ice ages occur, the effect has since been demonstrated in a plethora of laboratory experiments, and has also been proposed to be responsible for the way in which biological sensing mechanisms function to take advantage of inherent background noise. Writing on page 995 of this issue³, Robert Badzey and Pritiraj Mohanty demonstrate the effect in a doubly clamped mechanical beam of nanometre size. But this is more than just another stochastic resonance demo: the authors' device (besides belonging to the buzzword class 'nano') also bestrides the murky gap between classical and quantum physics.

Badzey and Mohanty's nanomechanical beam, $8\text{ }\mu\text{m} \times 200\text{ nm} \times 300\text{ nm}$ in size, is made of crystalline silicon of high purity, clamped at either end. When small-amplitude radiofrequency excitations are applied to it, the beam behaves like a damped, driven harmonic oscillator, showing a standard, 'lorentzian' shape (that is, peaked at a resonant frequency) to its response–power spectrum. But when the forcing amplitude exceeds a critical value, determined by (among other things) the amount of power dissipated as heat and the natural frequency at which the beam resonates, its response becomes nonlinear — the beam buckles.

This is not a new phenomenon: the concept of buckling dates back to the eighteenth century, when Leonhard Euler developed the partial differential equations that describe elastic instability. This 'Euler instability' dictates that a slender elastic object, clamped at either end and initially in stable equilibrium, will move transversely if subjected to a longitudinal compressive force. If this force is increased beyond a critical value, the object will undergo a transition to an unstable equilibrium point. (Ref. 4 provides a simple derivation of this critical force.)

In Badzey and Mohanty's system³, the compressive force comes from the application of a transverse driving force that changes the beam's length, making it longer than the distance between its two supports. When the beam vibrates from side to side, it is subjected

to an additional compressive force from the clamping pads as it passes through equilibrium. If the driving force is large enough, the beam responds nonlinearly and acquires two different vibrational modes at a single frequency near resonance. Crucially, these buckled modes are not static, but dynamic, and are underpinned by a 'bistable' potential-energy function; that is, one consisting of two stable states, or 'wells', separated by an energy threshold. The bistable configuration — together with an applied periodic signal whose amplitude is smaller than the threshold — is a basic condition for stochastic resonance.

Badzey and Mohanty apply noise to their system and compute a signal-to-noise ratio at the applied signal frequency for a series of response–power spectra, each taken at a different noise power. The resulting curve reveals a maximum signal-to-noise ratio at a critical noise power. This is because, when there is too little noise, it fails to lift the system over the threshold between the two states: the system cannot switch between them, and so there is no flow of information (assumed to occur in the switching events between wells). Conversely, too much noise leads to a surfeit of

(mostly incoherent) switching events and a corresponding loss of information. But between these two information minima, the switchings acquire maximum coherence with respect to the applied signal, and an information maximum occurs. Thus we have an intuitive result from otherwise counter-intuitive behaviour — that applying noise can be helpful.

The result³ conclusively demonstrates stochastic resonance in a nanomechanical oscillator at megahertz frequencies. But its significance does not end there. A basic criterion for the creation of a quantum harmonic oscillator is that the energy hf of the oscillator (where h is Planck's constant and f the oscillation frequency) must be larger than the broadening of the oscillator's energy spectrum that results from thermal effects. This energy spread is given by $k_B T$, where k_B is Boltzmann's constant and T is the temperature. At a temperature of 48 mK, for example, an oscillation frequency of 1 GHz ensures that this condition is met. Thus, the demonstration of stochastic resonance in the megahertz range paves the way for exploring the phenomenon in related structures — recently constructed by the same group^{5,6} — that have higher, quantum-regime frequencies.

Specifically, stochastic resonance might be used to enhance and control a quantum-mechanical signal in a nanomechanical oscillator — to achieve coherent control of oscillations between two quantum energy levels, for example. The ramifications of a 'quantum' stochastic resonance effect, and its possible relevance to low-temperature macroscopic quantum tunnelling, are fascinating.

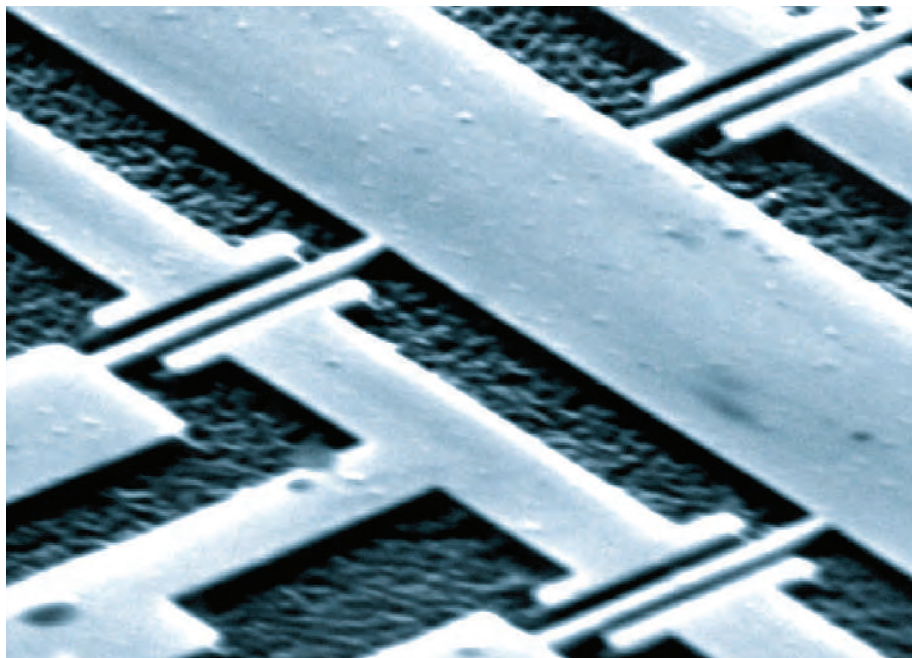


Figure 1 | Hang in there. A scanning electron micrograph of a set of suspended silicon nanomechanical beams with electrostatic control gates. Each beam can be made to move between two buckled states (0 or 1) by applying a field to its control gates. Stochastic resonance could allow enhanced control in switching between the two discrete states, and use of these nonlinear states as nanomechanical memory cells.

R. L. BADZEY & P. MOHANTY

They could pave the way to experiments in the as yet somewhat esoteric field of quantum measurement and control. For unsatisfied sceptics, Badzey and Mohanty have also developed a system of suspended nanomechanical beams made of silicon, each of which can be forced electrostatically to switch between its two stable states (Fig. 1). With stochastic resonance allowing enhanced control over such switching, these nonlinear states may eventually be useful as nanomechanical memory cells.

Stochastic resonance, although difficult to implement in practice, has always been an intriguing option for harnessing the background noise in certain systems. But the hubris of the 1990s was followed by a dearth of results to confirm that the mechanism underlies natural neurophysiological functions, and a general paucity of devices that were readily 'tunable' to take advantage of noise. These factors, together with the witches' brew of funding vicissitudes and some questionable publications, led to a waning of interest. Yet work such as that of Badzey and Mohanty³ shows that the effect can be invoked in a noisy, nonlinear dynamic system under appropriate operating conditions, and that it can also be exploited in carefully crafted applications. The

effect has also recently been used in biomedicine⁷ and in the amplification of electric field signals in carbon nanotube transistors⁸. These achievements, and the demonstration that background electrical noise is involved in the 'hunting frenzy' of the paddlefish *Polyodon spathula* as it preys on swarms of the water-flea *Daphnia*^{9,10}, indicate that this effect is more than just a laboratory curiosity. ■

Adi R. Bulsara is at the Space and Naval Warfare Systems Center, 53560 Hull Street, San Diego, California 92152-5001, USA, and is a visiting scientist at the US Office of Naval Research Global, London, UK.
e-mail: bulsara@spawar.navy.mil

1. Wiesenfeld, K. & Moss, F. *Nature* **373**, 33–36 (1995).
2. Bulsara, A. & Gammaitoni, L. *Phys. Today* **49**(3), 39–45 (1996).
3. Badzey, R. L. & Mohanty, P. *Nature* **437**, 995–998 (2005).
4. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. 1, Ch. 38 (Addison-Wesley, Redwood, CA, 1989).
5. Gaidarzhy, A. et al. *Phys. Rev. Lett.* **94**, 030402 (2005).
6. Gaidarzhy, A. et al. *Appl. Phys. Lett.* **86**, 254103 (2005).
7. Priplata, A., Niemi, J., Henry, J., Lipsitz, L. & Collins, J. *Lancet* **362**, 1123–1124 (2003).
8. Lee, I., Liu, X., Kosko, B. & Zhou, C. *Nano Lett.* **3**, 1683–1686 (2003).
9. Russell, D., Wilkens, L. & Moss, F. *Nature* **402**, 291–294 (1999).
10. Collins, J. *Nature* **402**, 241–242 (1999).

DEVELOPMENTAL BIOLOGY

Cell cycle unleashed

Takeo Kishimoto

How does fertilization cause animal eggs to begin embryonic development? Following entry of the sperm, the ingeniously regulated degradation of a protein seems to kick-start the stalled cell cycle.

In animal eggs, the cell-division cycle is held in check part-way through, awaiting sperm entry. A major event after fertilization is therefore the alleviation of this blockage so that cell division can begin in earnest to form the embryo. Thomas Mayer and colleagues (page 1048 of this issue)¹ and Liu and Maller, writing in *Current Biology*², now provide a molecular answer to the long-standing question of how the cell-cycle arrest is released by fertilization.

In sexual reproduction, development of the embryo cannot begin until completion of the specialized cell cycle that forms eggs (meiosis). This temporal coupling is coordinated such that the meiotic cell cycle in eggs is arrested at a particular stage and the arrest is released by fertilization. In vertebrates, including frogs, mice and humans, the stage of the cycle at which the cell typically arrests is called metaphase of meiosis II (or meta-II)³. Relieving the cell-cycle blockage requires the activity of a protein complex called APC/C^{Cdc20}. This complex is a highly regulated enzyme that

targets several cell-cycle regulatory proteins for destruction by attaching a ubiquitin group to them. Removal of these regulatory proteins then allows the cell to exit from metaphase and move on to the next stage of the cell cycle⁴.

Early studies in the 1930s by Lewis Victor Heilbrunn and Daniel Mazia showed that one of the earliest molecular events following fertilization is a rapid escalation in intracellular calcium ion concentration, and they suggested that this increase might be the signal that triggers embryo development. In frog and mouse eggs, this rise in calcium activates a protein called calmodulin-dependent protein kinase II (CaMKII)⁵. But how the activity of APC/C^{Cdc20} is inhibited during meta-II arrest, and how CaMKII abolishes that inhibition, have been largely unknown. Using extracts from frog eggs, Mayer and colleagues¹ and Liu and Maller² demonstrate that CaMKII acts on the protein Erp1 (for 'Emi1-related protein 1'), an inhibitor of APC/C^{Cdc20}. This causes Erp1 to be degraded and thereby allows APC/C^{Cdc20} to release the brakes on the cell cycle (Fig. 1).



50 YEARS AGO

In the leading article in *Nature* of August 20 on "Educational Problems of the Colonial Territories", it is stated that "only some 450 scientists are at present engaged in Colonial research"... Most British scientific workers are superannuated at the age of approximately sixty-five. Many of them are capable of another ten years of research, and a moderate amount of teaching. Some, at least, would be happy to work in a Colonial university or research institute. The necessary qualifications are the capacities to work in a tropical or subtropical climate, and to form friendships with non-Europeans... I think that the presence in a Colonial university of even two or three Fellows of the Royal Society carrying out fundamental research with African or Asian colleagues would help the local population to generate its own scientific culture.

J. B. S. Haldane

From *Nature* 15 October 1955.

100 YEARS AGO

The Citizen, a Study of the Individual and the Government — Prof. Shaler, who is professor of geology at Harvard, has set before himself the practical and unambitious task of instructing the youth of the United States in the first principles of citizenship. In this he has succeeded; his work is interesting, suggestive and extremely sensible... A favourable specimen of his mode of argument may be found in the discussion of woman's suffrage. There is no reference to the various views held by thinkers from Plato downwards; but probably Prof. Shaler's one-page argument is quite sufficient, that women, owing to their usually secluded lives, are not fitted in the same way as men to form judgments on political questions, but that, after all, if a majority of women should desire to vote, it would probably be best to give them the franchise, for the reason that it is most undesirable to have any considerable body of the people in a discontented state.

From *Nature* 12 October 1905.

50 & 100 YEARS AGO

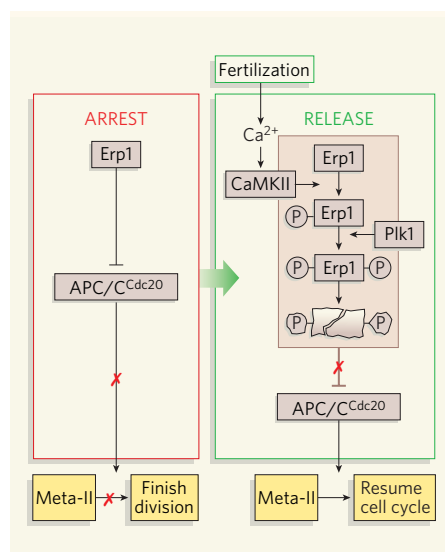


Figure 1 | Releasing the brakes. Frog eggs awaiting fertilization are arrested at the cell-division stage called metaphase of meiosis II (meta-II), owing to the inhibition of the APC/C^{Cdc20} protein complex. This complex acts on certain inhibitors of the next stage in the cell cycle (anaphase), targeting them for degradation. The inhibition of APC/C^{Cdc20} requires Erp1. Upon fertilization, the resulting Ca²⁺ rise activates calmodulin-dependent protein kinase II (CaMKII). Mayer and colleagues¹ show that CaMKII phosphorylates Erp1 (circled P), thereby creating a docking site for Polo-like kinase 1 (Plk1). The recruited Plk1 phosphorylates Erp1 again, thereby targeting it for destruction. As a result, APC/C^{Cdc20} is no longer inhibited, and anaphase inhibitors such as mitotic cyclins and securin are targeted for degradation, leading to the release of meta-II arrest. Thus, CaMKII acts as a novel priming kinase for Erp1 phosphorylation by Plk1, which links the fertilization signal to resumption of the cell cycle. Red crosses indicate steps in the pathway that do not occur.

Mayer and colleagues¹ further clarify the mechanisms by which CaMKII targets Erp1 for destruction, building on three previous findings concerning a protein called Polo-like kinase 1 (Plk1). This protein choreographs several steps in meiosis and mitosis — the basic cell-division cycle that produces most cells in the body⁶. First, Plk1 is required for APC/C^{Cdc20} activation at the release from meta-II arrest⁷. Second, Erp1 is identified as a novel APC/C^{Cdc20} inhibitor that is required for meta-II arrest, and it can be phosphorylated by Plk1 and thereby targeted for degradation⁸. Third, Plk1 contains a motif called the Polo-box domain (PBD) that can bind to a specific peptide 'docking' sequence in another protein, but only once the peptide has been phosphorylated by a 'priming kinase'. This suggests a mechanism whereby the priming kinase phosphorylates the PBD-docking motif in Plk1 substrates, allowing Plk1 to bind to the substrate and phosphorylate it a second time^{6,9}.

Mayer and colleagues¹ demonstrate that CaMKII functions as a priming kinase for

Erp1 — allowing it then to be recognized and phosphorylated by Plk1. The second phosphorylation of Erp1 targets it for degradation, releasing APC/C^{Cdc20} from inhibition. This then triggers the exit from meta-II and resumption of the cell cycle. So the calcium increase after fertilization sets in motion a precisely regulated system of protein degradation that eventually releases meta-II arrest (Fig. 1). The same pathway probably functions in eggs of other vertebrates. This model explains why Erp1 is not attacked by Plk1 until fertilization, even though Plk1 is active during meta-II arrest.

A close relative of Erp1, called Emi1 (early mitotic inhibitor 1), is also found in frog eggs and can also inhibit APC/C^{Cdc20} (ref. 10). Although Emi1 was initially thought to be essential for meta-II arrest in the same way as Erp1 (ref. 11), this possibility was excluded by its absence during meta-II arrest¹². Whether Erp1 is involved in the mitotic cell cycles in the early embryo is not known, but it is plausible that Erp1 inhibits APC/C^{Cdc20} in the egg and early embryo, and that Emi1 takes over this role later in development. Jackson and colleagues¹⁰ propose that Emi1 inhibits the APC/C^{Cdc20} complex by binding to the Cdc20 protein, and it is possible that Erp1 acts in a similar manner.

Although the studies of Mayer and colleagues^{1,8} and of Liu and Maller² have unravelled the mechanism by which meta-II arrest is relieved following fertilization, it remains unclear how arrest occurs in the first place. More than 30 years ago, Masui and Markert identified cytotstatic factor (CSF), a cytoplasmic activity responsible for frog meta-II arrest³. Further studies established that CSF

activity includes a signalling pathway that involves the Mos, MAPK and Rsk proteins (reviewed in refs 3, 13). Although another proposed initiator of meta-II arrest is the spindle checkpoint pathway¹³, it seems that Mad2, a core component of this pathway, may not be essential for CSF activity². But if Erp1 can itself inhibit APC/C^{Cdc20}, why is the Mos–MAPK–Rsk pathway required for meta-II arrest? Presumably, Erp1 can cooperate with Mos–MAPK–Rsk to prevent Cdc20 from activating APC/C^{Cdc20}. If so, CSF may consist of both Erp1 and Mos–MAPK–Rsk. This question bears further investigation: Erp1 may turn out to be key to the arrest of the cell cycle while the egg awaits fertilization, as well as to its subsequent resumption after sperm entry.

Takeo Kishimoto is in the Laboratory of Cell and Developmental Biology, Graduate School of Bioscience, Tokyo Institute of Technology, Nagatsuta, Midoriku, Yokohama 226-8501, Japan.

e-mail: tkishimo@bio.titech.ac.jp

1. Rauh, N. R., Schmidt, A., Bormann, J., Nigg, E. A. & Mayer, T. U. *Nature* **437**, 1048–1052 (2005).
2. Liu, J. & Maller, J. L. *Curr. Biol.* **15**, 1458–1468 (2005).
3. Masui, Y. *Nature Rev. Mol. Cell Biol.* **1**, 228–232 (2000).
4. Peters, J. M. *Mol. Cell* **9**, 931–943 (2002).
5. Lorca, T. et al. *Nature* **366**, 270–273 (1993).
6. Barr, F. A., Sillje, H. H. W. & Nigg, E. A. *Nature Rev. Mol. Cell Biol.* **5**, 429–440 (2004).
7. Descombes, P. & Nigg, E. A. *EMBO J.* **17**, 1328–1335 (1998).
8. Schmidt, A. et al. *Genes Dev.* **19**, 502–513 (2005).
9. Elia, A. E. H. et al. *Cell* **115**, 83–95 (2003).
10. Reimann, J. D. R. et al. *Cell* **105**, 645–655 (2001).
11. Reimann, J. D. R. & Jackson, P. K. *Nature* **416**, 850–854 (2002).
12. Ohsumi, K., Koyanagi, A., Yamamoto, T. M., Gotoh, T. & Kishimoto, T. *Proc. Natl Acad. Sci. USA* **101**, 12531–12536 (2004).
13. Tunquist, B. J. & Maller, J. L. *Genes Dev.* **17**, 683–710 (2003).

PLANT PHYSIOLOGY

A big issue for trees

Josep Peñuelas

The age of a tree and its size tend to increase together. Disentangling the effects of these two factors on tree vitality is no easy task, but further evidence adds to the view that it is size that matters.

What controls senescence and lifespan? As they describe in *Ecology Letters*, Mencuccini *et al.*¹ have tackled this question for trees, which may live for many centuries and grow to more than 100 metres in height. With increasing age and size, growth tends to slow and a tree is more likely to die. One explanation for this is that reduced growth results from tissue and cell senescence, which is under direct genetic or age-related control. Another is the physiological burdens, such as the demands on water and nutrient supply, that are associated with increasing height and girth. Mencuccini and colleagues' results add experimental support for this second view.

The term senescence encompasses a collection of changes that are associated with ageing. Such changes are triggered by altered gene expression, and include a progressive loss of physiological functions, a decrease of fertility and a greater vulnerability to disease or damage. It has long been believed that senescence is an inevitable consequence of ageing in all plants and animals. Evidence from diverse disciplines has challenged this assumption for animals², however, and there are indications that it may also not apply to trees.

First, there is a difference between cellular, tissue or organ senescence and whole-tree senescence. The former process is fairly well



Figure 1 | Lonesome pine. This example of an ancient bristlecone pine stands in the Great Basin National Park, Nevada.

known: for example, leaf senescence occurs as a process of programmed cell death³, and is age-dependent and induced by environmental cues such as extremes in temperature, light intensity and duration, or moisture⁴. But there is little evidence of changes in gene expression in ageing trees⁵, and therefore of genetically controlled mechanisms to explain the reduced growth of mature specimens.

Second, the reproductive period of trees can extend over centuries. Their reproductive output, instead of declining, increases with age (and therefore size) and the number of potential reproductive buds.

Third, woody plants such as the famous Cabernet Sauvignon grapevines have been propagated for more than 800 years by serial grafting⁶, and aspen clones can apparently replicate indefinitely. So it seems that at least some cell lines inside the meristems — the growing points from which new cells form — retain the juvenile ability to contribute to new growth even in ageing stock.

Finally, slow-growing non-clonal trees can live for millennia, and apparently still be in good shape. Such trees tend to inhabit extreme

environments. For example, bristlecone pines atop the arid mountains in the western United States are more than 4,000 years old (Fig. 1). Trees on the vertical surfaces of cliffs likewise endure harsh conditions, grow only slowly and may live for more than 1,000 years⁷. Slow growth minimizes maintenance and repair costs, while maximizing durability and strength.

With this as background, Mencuccini *et al.*¹ carried out an experimental test of the presence or absence of senescence-related changes in trees. Their study involved four tree species representing different evolutionary groups (gymnosperms and angiosperms), and with different water-transport characteristics (ring-porous, diffusive-porous and tracheid-bearing), leaf ecology (evergreen and deciduous) and intensity of management practices (unmanaged and intensively managed). For each species, they measured growth, together with the gas-exchange and biochemical properties of leaves, in trees of different ages (1–269 years old) and sizes (2–42 metres in height) in the field. They repeated the measurements on specimens — now of equally small size — that

had been propagated, by grafting or by direct rooting, from the same trees.

Growth, net photosynthetic rates and leaf nitrogen concentrations declined in the field with size and age, but there were no such declines in the corresponding propagated plants. Mencuccini and colleagues conclude, therefore, that size, not meristematic cellular senescence, accounts for the reduced growth rates in older, taller trees.

As the authors themselves point out, however, it is possible that factors or conditions elsewhere in the tree may still trigger senescence in the meristem. These system-level signals could have disappeared in the small grafted plants, with the meristems reverting to a more juvenile condition. To check this possibility, the same measurements need to be conducted on stocks of different sizes. Such work has in fact been done on Japanese cedars⁸, in which it was found that grafted shoots in the upper crowns of tall cedars showed a similarly poor performance to intact shoots in those crowns. These results provide a further indication that extrinsic factors mediated by size, not irreversible intrinsic changes in the meristems, drive the decline in photosynthetic rates in larger trees.

Given these observations, we might wonder whether, if trees remain short, they will remain vigorous. Bonsais can live for hundreds of years: do they never senesce, and do they live longer than their big brothers? And do the short trees characteristic of Mediterranean climates live longer than tall trees in moist temperate regions? It still seems unlikely that size alone determines tree vigour.

The way forward lies in combining research in plant physiological ecology and molecular biology, carried out on both trees and grafts. With more cross-talk between these disciplines, the aim will be to identify possible biochemical and genetic indicators of senescence in meristematic and non-meristematic tissues, and to relate those findings to constraints imposed by the tree as a whole. The outcome will be of more than academic value, and of keen interest, for example, to those involved in forest conservation and timber production. ■ Josep Peñuelas is at the Consejo Superior de Investigaciones Científicas (CSIC-CEAB) and the Center for Ecological Research and Forestry Applications (CREAF), Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain. e-mail: josep.penuelas@uab.es

1. Mencuccini, M. *et al.* *Ecol. Lett.* doi:10.1111/j.1461-0248.2005.00819.x (2005).
2. Helfand, S. L. & Inouye, S. *Nature Rev. Genet.* **3**, 149–153 (2002).
3. Yoshida, S. *Curr. Opin. Plant Biol.* **6**, 79–84 (2003).
4. Munne-Bosch, S. & Peñuelas, J. *Ann. Bot.* **92**, 385–391 (2003).
5. Diego, L. B. *et al.* *J. Exp. Bot.* **55**, 1597–1599 (2004).
6. Noodén, L. D. & Thompson, J. E. in *Handbook of the Biology of Aging*, 2nd edn (eds Finch, C. E. & Schneider, E. L.) 105–127 (Van Nostrand, New York, 1985).
7. Larson, D. W. *et al.* *Nature* **398**, 382–383 (1999).
8. Matsuzaki, J., Norisada, M., Kodaira, J., Suzuki, M. & Tange, T. *Trees* **19**, 198–203 (2005).

BRIEF COMMUNICATIONS

Millet noodles in Late Neolithic China

A remarkable find allows the reconstruction of the earliest recorded preparation of noodles.

Noodles have been a popular staple food in many parts of the world for at least 2,000 years¹, although it is debatable whether the Chinese, the Italians or the Arabs invented them first. Here we analyse a prehistoric sample of noodles contained in a well preserved, sealed earthenware bowl discovered in the Late Neolithic^{2–4} archaeological site of Lajia in northwestern China. We identify millet as the source of the abundant seed-husk phytoliths and starch grains present in the vessel. This shows that the conversion of ground millet flour into dough that could be repeatedly stretched into long, thin strands for the preparation of boiled noodles was already established in this region 4,000 years ago.

The Lajia archaeological site (35° 49' 40" N, 102° 51' 15" E) is located on a terrace on the upper reaches of the Yellow River in northwestern China, and has been excavated since 1999 (refs 2, 3). The Neolithic cultural settlement² containing the prehistoric bowl of noodles was found beneath a floodplain sediment layer that was about 3 metres thick. Radiocarbon (¹⁴C) measurements date its occupation to around 4,000 yr BP (ref. 4) (see supplementary information). A very large earthquake and catastrophic flooding probably destroyed the Lajia settlement at about this time³.

The bowl was found upside-down and embedded in brownish-yellow, fine clay. When we lifted off the bowl, the remains of the noodles were found inside, on top of the cone of sediment that had filled the inverted earthenware container (Fig. 1a). The noodles were thin (about 0.3 cm in diameter), delicate, more than 50 cm in length and yellow in colour. They resemble the La-Mian noodle, a traditional Chinese noodle that is made by repeatedly pulling and stretching the dough by hand.

To determine the taxa of the cereals present at the site at the time, and to establish which might have been used for the fine flour needed for noodle preparation, we analysed the phytoliths and starch grains present in the sediment associated with the noodles. Assignment of phytolith morphotypes and starch grains to plant taxa was based on our modern reference collection of more than 85 grasses, commercial plants and weeds native to the study region. In addition, we used published keys of phytoliths and starch grains to aid identification^{5–9}.

We paid particular attention to plant taxa cultivated in northwest China^{10–12}. Our inves-

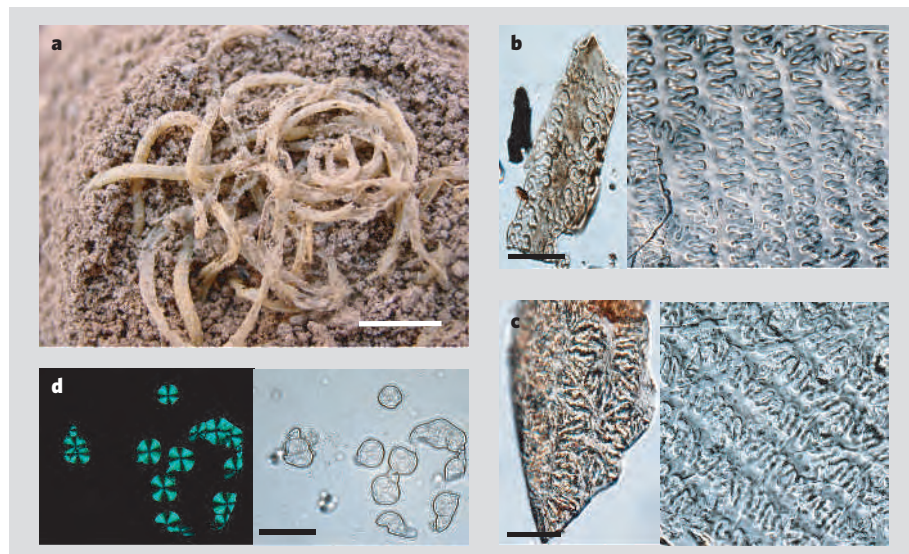


Figure 1 | Late Neolithic noodles from China. **a**, Noodles dating to 4,000 years ago, shown here on top of an in-filled sediment cone and revealed after the inverted earthenware bowl containing them was removed. Scale bar, 1 cm. **b**, Dendriform-1 husk phytoliths (left) from the noodle sample compared with husks from the modern millet *Panicum miliaceum* (right). **c**, Dendriform-2 husk phytoliths from the noodles (left) compared with husks from the modern millet *Setaria italica* (right); the dendriform-2 type mostly occurred at the lemma end of *S. italica*. **d**, Polarized-light (left) and light micrographs (right) of starch grains from the prehistoric noodles. Although the lamellae characteristics of noodle starch were mostly lost as a result of gelatinization during cooking, their size and cross-shaped birefringence under polarized light are similar to those of starch from the millets *S. italica* and *P. miliaceum*. Scale bar, 20 μ m for **b**, **c**, and **d**.

tigations revealed that the genera *Hordeum* (barley), *Triticum* (wheat), *Panicum* (broom corn millet) and *Setaria* (foxtail millet) were distinguishable from other grasses native to the study region by their phytolith patterning and by the morphology and size of their starch grains.

We examined six phytolith samples from the sediment that had filled the bowl, and used them to diagnose their plant of origin. Three samples from the noodle layer near the bottom of the bowl contained abundant husk phytoliths of the types known as dendriform-1 (Fig. 1b) and dendriform-2 (Fig. 1c) at concentrations of 9.81×10^4 and 4.36×10^4 grains per gram, respectively. Their shape and patterning allowed us to diagnose them as millets, and they were identified as belonging to *Panicum miliaceum* and *Setaria italica*, respectively (see supplementary information). The starch grains found in the noodle sample (Fig. 1d) closely resemble those of millets, based on our comparison with modern reference samples from different cereal grains.

Unlike modern Italian pasta and Asian

noodles, which are generally made from durum wheat (tetraploid) and bread wheat (hexaploid), respectively, the prehistoric noodles show no evidence that wheat, barley or other non-grass plants were used to supply their ingredients. Our findings support the belief that early plant domestication and food production relied on millet crops^{10–12} in the semi-arid Loess Plateau region of China.

Houyuan Lu[†], Xiaoyan Yang^{*}, Maolin Ye[‡], Kam-Biu Liu[§], Zhengkai Xia[¶], Xiaoyan Ren[#], Linhai Cai[#], Naiqin Wu^{*}, Tung-Sheng Liu^{*}

^{*}Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China
e-mail: houyuanlu@mail.iggcas.ac.cn

[†]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China

[‡]Institute of Archaeology, Chinese Academy of Social Sciences, Beijing 100710, China

[§]Department of Geography and Anthropology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

[¶]College of Environmental Science, Peking University, Beijing 100871, China

[#]Qinghai Provincial Institute of Cultural Relics

and Archaeology, Sining 810007, China

- Hou, G. *Adv. Food Nutr. Res.* **43**, 143–193 (2001).
- Institute of Archaeology, Chinese Academy of Social Sciences & Qinghai Provincial Institute of Cultural Relics and Archaeology *Archaeology* **12**, 12–25 (2002).
- Yang, X. Y., Xia, Z. K. & Yi, M. L. *Chinese Sci. Bull.* **48**, 1877–1881 (2003).
- Institute of Archaeology *Archaeology* **12**, 58–76 (2003).
- Ball, T. B., Gardner, J. S. & Anderson, N. *Am. J. Bot.* **86**, 1615–1623 (1999).
- Pearsall, D. M. *Paleoethnobotany: A Handbook of Procedures* 2nd edn (Academic, San Diego, 2000).
- Lu, H. Y. & Liu, K. B. *Div. Distrib.* **9**, 73–87 (2003).
- Piperno, D. R., Weiss, E., Holst, I. & Nadel, D. *Nature* **430**, 670–673 (2004).
- Fujita, S., Sugimoto, Y., Yamashita Y. & Fuwa, H. *Food Chem.* **55**, 209–213 (1996).
- Shelach, G. J. *World Prehist.* **14**, 363–413 (2000).
- Institute of Archaeology, Chinese Academy of Social Sciences & Qinghai Provincial Institute of Cultural Relics and Archaeology *Archaeol. Cult. Relics* **2**, 85–91 (2004).
- Ren, S. N. *Archaeology* **1**, 37–49 (1995).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

doi:10.1038/437967a

MICROWAVE DEVICES

Carbon nanotubes as cold cathodes

To communicate, spacecraft and satellites rely on microwave devices, which at present are based on relatively inefficient thermionic electron sources that require heating and cannot be switched on instantaneously. Here we describe a microwave diode that uses a cold-cathode electron source consisting of carbon nanotubes¹ and that operates at high frequency and at high current densities.

Because it weighs little, responds instantaneously and has no need of heating, this miniaturized electron source should prove valuable for microwave devices used in telecommunications.

We constructed a microwave diode in which the carbon-nanotube field-emission source was directly driven at gigahertz (GHz) frequencies. Arrays of vertically aligned carbon nanotubes were integrated on a coaxial post in a resonant cavity. In the device simulation shown in Fig. 1a, radiofrequency electromagnetic radiation at the input induces a high, oscillating electric field at the end of the coaxial post; this electric field is further amplified by the carbon-nanotube array (for details, see supplementary information).

The carbon-nanotube array (Fig. 1b) consists of uniform individual carbon nanotubes² spaced at a distance corresponding to roughly twice their height in order to minimize electrostatic-field shielding from adjacent emitters³. Each cathode has an active area of $0.5 \times 0.5 \text{ mm}^2$ (or 2,500 carbon nanotubes) and 16 cathodes can be created simultaneously (Fig. 1b, inset).

The device was operated at 1.5 GHz using various radiofrequency-input powers to generate different macroscopic electric fields at the array of carbon-nanotube emitters. As the cavity walls and emitters are grounded, the radiofrequency electric field exists only inside the cavity, as shown in the equivalent electrical circuit in Fig. 1c. A spectrum analyser connected to the output antenna confirmed the presence of the fundamental 1.5-GHz peak in the cavity. In this study, cathodes were

operated at 1 mA and 1.5 GHz for 40 h without degradation or a decrease in current output (within the measurement error of 5%). With an applied radiofrequency electric field of 29 megavolts per metre, the output at the anode reaches 3.2 mA, with an average current density of 1.3 A cm^{-2} (Fig. 1d). This corresponds to a peak current of 30 mA and a current density of 12 A cm^{-2} in the output waveform (see supplementary information).

The ability directly to generate or modulate an electron beam at high current density and gigahertz frequencies from carbon nanotubes is an important technological advance. Thermionic sources used in today's microwave devices are operated by direct current or at low frequency; their electron beam is usually modulated downstream in an extended interaction line, leading to physically long devices. In contrast, carbon-nanotube cold cathodes that have a vacuum gap to a stand-off grid or anode of a few hundred micrometres or less, as we describe here, have low capacitances and can be operated at very high frequencies (for example, 32-GHz modulation of carbon-nanotube emitters has been achieved from a microwave diode and triode; L.H. *et al.*, manuscript in preparation). They can therefore be used directly as the input stage of a microwave amplifier.

Carbon-nanotube emitters are robust and do not suffer from electromigration because of their strong C–C covalent bonding. Metal emitters, on the other hand, often fail owing to field-induced sharpening, which leads to thermal runaway of the emitters.

Our carbon-nanotube cathode already delivers average- and peak-current densities that are similar to those used in present-day microwave transmission devices. Because of their small size, and ability to generate and modulate the beam directly and on demand without the need for high temperatures, carbon-nanotube cathodes hold promise for a new generation of lightweight, efficient and compact microwave devices for telecommunications in satellites or spacecraft.

Kenneth B. K. Teo*, **Eric Minoux†**, **Ludovic Hudanski‡**, **Franck Peauger‡**, **Jean-Philippe Schnell†**, **Laurent Gangloff***, **Pierre Legagneux†**, **Dominique Dieumegard‡**, **Gehan A. J. Amaratunga***, **William I. Milne***

*Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK
e-mail: wim@eng.cam.ac.uk

†Thales Research and Technology, RD 128, 91767 Palaiseau Cedex, France

‡Thales Electron Devices, BP 23, 78141 Vélizy-Villacoublay, France

- de Jonge, N. & Bonard, J. M. *Phil. Trans. R. Soc. Lond. A* **362**, 2239–2266 (2004).
- Teo, K. B. K. *et al.* *Nanotechnology* **14**, 204–211 (2003).
- Nilsson, L. *et al.* *Appl. Phys. Lett.* **76**, 2071–2073 (2000).

Supplementary information accompanies this communication on Nature's website

Competing financial interests: declared none.

doi:10.1038/437968a

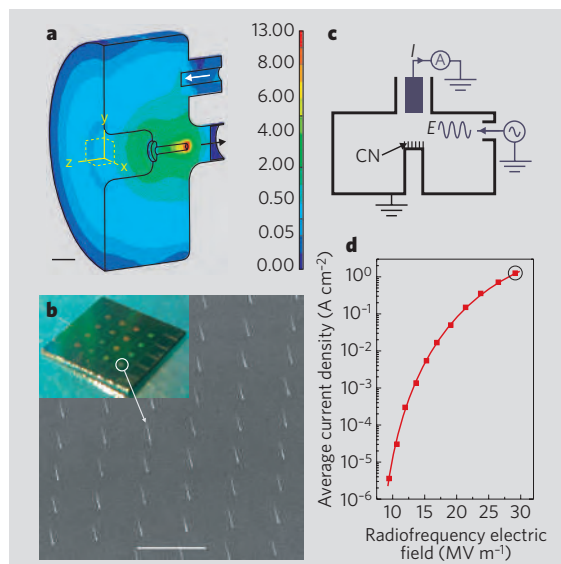


Figure 1 | Features of the carbon-nanotube microwave diode. **a**, Simulation of the coaxial resonant cavity (a cross-section is shown) that was used to generate a high electric field (red) at the carbon-nanotube-array cathode from the radiofrequency input; colour scale shows the applied macroscopic electric field in volts ($\times 10^5$) per metre. White arrow, coaxial radiofrequency input; black arrow, emitted electron beam, collected by an antenna; scale bar 10 mm. **b**, Electron micrograph of the carbon-nanotube-array cold cathode at a tilt of 45° . The carbon nanotubes have an average diameter of 49 nm, height of 5.5 μm and a spacing of 10 μm ; scale bar, 15 μm . Inset, photograph of 16 cathodes. **c**, Representation of the equivalent electrical circuit, where E is the applied electric field and I is the emitted current; CN, carbon nanotube array. **d**, Measured average current density plotted against applied radiofrequency electric field using 1.5-GHz sinusoidal input. The circled point corresponds to $I = 3.2 \text{ mA}$. The cavity-quality factor was 3,160 (see supplementary information).

Implications for prediction and hazard assessment from the 2004 Parkfield earthquake

W. H. Bakun¹, B. Aagaard¹, B. Dost², W. L. Ellsworth¹, J. L. Hardebeck¹, R. A. Harris¹, C. Ji³, M. J. S. Johnston¹, J. Langbein¹, J. J. Lienkaemper¹, A. J. Michael¹, J. R. Murray¹, R. M. Nadeau⁴, P. A. Reasenberg¹, M. S. Reichle⁵, E. A. Roeloffs⁶, A. Shakal⁵, R. W. Simpson¹ & F. Waldhauser⁷

Obtaining high-quality measurements close to a large earthquake is not easy: one has to be in the right place at the right time with the right instruments. Such a convergence happened, for the first time, when the 28 September 2004 Parkfield, California, earthquake occurred on the San Andreas fault in the middle of a dense network of instruments designed to record it. The resulting data reveal aspects of the earthquake process never before seen. Here we show what these data, when combined with data from earlier Parkfield earthquakes, tell us about earthquake physics and earthquake prediction. The 2004 Parkfield earthquake, with its lack of obvious precursors, demonstrates that reliable short-term earthquake prediction still is not achievable. To reduce the societal impact of earthquakes now, we should focus on developing the next generation of models that can provide better predictions of the strength and location of damaging ground shaking.

Earthquake prediction is the Holy Grail of seismology. Although the ability to predict the time and location of earthquakes remains elusive, predicting their effects, such as the strength and geographical distribution of shaking, is routine practice. The extent to which earthquake phenomena can accurately be predicted will ultimately depend on how well the underlying physical conditions and processes are understood. To understand earthquakes requires observing them up close and in detail—a difficult task because they are at present largely unpredictable, and so knowing where to put the instrumentation needed to make such observations is a challenge. The 40-km-long Parkfield section of the San Andreas fault was recognized two decades ago as a promising earthquake physics laboratory and an intensive experiment was established to record the next segment-rupturing earthquake there and provide the much-needed detailed observations. The occurrence of the anticipated moment magnitude $M_w = 6.0$ earthquake on 28 September 2004 (origin time 17:15:24 Coordinated Universal Time, UTC; epicentre location 35.815° N, 120.374° W; depth 7.9 km) fulfilled that promise.

The Parkfield section of the San Andreas fault is bounded on the northwest by a 150-km-long creeping section, where numerous small earthquakes occur, and on the southeast by hundreds of kilometres of locked fault where few earthquakes have been detected in the twentieth century (Fig. 1). The 1857 $M_w = 7.9$ Fort Tejon earthquake ruptured the locked fault southeast of Parkfield and is thought to have initiated near Parkfield¹. On the Parkfield section, the motion of the Pacific plate relative to the North America plate is partly accommodated by repeating $M_w = 6.0$ earthquakes. The historical record of earthquakes at Parkfield includes at least six such events since 1857 (ref. 2; Supplementary Text 1, Supplementary Fig. 1 and Supplementary Table 1).

The simple setting and apparent regularity of Parkfield earthquakes³ offered the rationale for the only scientific earthquake prediction officially recognized by the United States government⁴ and an opportunity to place instruments in the region before the anticipated earthquake. The primary goal of the Parkfield Earthquake Prediction Experiment⁵ was to obtain a detailed understanding of the processes leading up to the anticipated earthquake; a secondary goal was to issue a public warning shortly before the earthquake. A variety of sensors were deployed in a dense network designed specifically to record the build-up of strain in the surrounding crust, monitor earthquakes and slip on the fault, and detect any precursors that might foreshadow a large earthquake. Complementary arrays of strong-ground-motion sensors were deployed to record shaking near the earthquake rupture zone⁶.

A significant development of the Parkfield experiment has been the collaboration of federal, state and local officials to develop a protocol for issuing short-term earthquake alerts⁷; the protocol provided a template for communication between scientists and emergency responders and subsequently served as a prototype for volcanic hazard warning protocols⁸. Innovations in the collection, transmission and storage of Parkfield data included a pioneering effort to provide publicly available, near-real-time earth science data streams over the internet. The systems pioneered at Parkfield have become standard elements of seismic monitoring throughout the US and have set the foundation for the installation of the USGS Advanced National Seismic System (ANSS)⁹. The Parkfield dense instrumentation network, which includes a variety of geophysical sensors, motivated the placement of a scientific borehole, Earthscope's San Andreas Fault Observatory at Depth (SAFOD), at the northwestern end of the Parkfield segment^{10,11} and served as a

¹US Geological Survey, Menlo Park, California 94025, USA. ²Royal Netherlands Meteorological Institute (KNMI), Seismology Division, PO Box 201, 3730AE De Bilt, The Netherlands. ³Division of Geological and Planetary Sciences, Caltech, Pasadena, California 91125, USA. ⁴University of California, Berkeley Seismological Laboratory, Berkeley, California 94720, USA. ⁵California Geological Survey, Sacramento, California 95814, USA. ⁶US Geological Survey, Vancouver, Washington 98683, USA. ⁷Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA.

prototype for the geodetic networks that are part of Earthscope's Plate Boundary Observatory¹¹.

In 1985, the USGS issued a long-term prediction that an earthquake of approximately $M_w = 6$ would occur before 1993 on the San Andreas fault near Parkfield⁴. After the prediction window closed, *sans* earthquake, an independent evaluation of the Parkfield Earthquake Prediction Experiment was conducted by a Working Group of the National Earthquake Prediction Evaluation Council¹². This working group recommended that monitoring be continued as a long-term effort to record the next earthquake at Parkfield. The failure of the long-term prediction of the time of the earthquake as well as certain aspects of the 2004 Parkfield earthquake (discussed below), have confounded nearly all simple earthquake models. However, the experiment's primary goal of observing a large earthquake near the rupture has been achieved. The instruments at Parkfield continue to operate and provide important data on post-seismic deformation and other processes. Altogether, these recordings are providing a picture, at unprecedented resolution, of what occurred before, during and after the 2004 earthquake^{13,14}. Although the analysis is far from complete, it is clear even now that the observations have important implications for nearly all areas of seismic hazard analysis and loss reduction.

Precursors and earthquake prediction

The idea that detectable precursory processes precede earthquakes dates back to at least the seventeenth century¹⁵. However, with the exception of foreshocks, unambiguous and repeatable instrumental observations of such phenomena remain elusive. As noted, identical

foreshocks preceded both the 1934 and 1966 Parkfield earthquakes by 17 min (ref. 16). However, such foreshocks did not precede the 1901 or 1922 events and so the Parkfield prediction experiment, which was designed to record potential foreshocks as well as other precursory signals, treated all precursors in a probabilistic manner⁷. At present, with the exception of an ambiguous low-level strain of $<10^{-8}$ that occurred during the 24 h before the main shock, there is no evidence of any short-term precursory signal, either seismic or aseismic¹³. Even microseismicity, detectable at the $M = 0$ threshold in the epicentral region, was absent during the six days before the main shock¹³. Notable precursory signals are not evident in the magnetic field, telluric electric field, apparent resistivity, or creep observations¹³. This lack of short-term precursors emphasizes the difficulty of reliable short-term earthquake prediction (up to a few weeks before).

Subtle strain changes of a few nanostrain were recorded on several instruments in the 24 h before the earthquake. Such changes place important constraints on earthquake nucleation physics but are too small to provide a reliable basis for issuing public warnings that a damaging earthquake is imminent. The dense instrumentation arrays continue to uncover new processes, such as deep tremor under the locked section of the fault southeast of the Parkfield segment¹⁷. Future hopes for prediction will rest on whether such processes are precursory or simply commonplace.

Fault structure and segment boundaries

The similar magnitude and rupture extent of the last six Parkfield earthquakes supports the concept of fault segmentation and the role of segment boundaries in influencing the rupture extent and magnitude of earthquakes. The nature of segment boundaries, however, is controversial. Fault geometry, rheological and frictional properties of materials, pore fluids and stress conditions have all been proposed to explain segment boundaries¹⁸.

Lindh and Boore¹⁹ suggested a fault-geometry-based explanation for the location of the boundaries of the Parkfield segment: a 5° bend in the fault trace to the northwest and a right-stepover to the southeast appeared to offer geometric obstacles that could limit earthquake rupture. The 2004 aftershocks relocated with a three-dimensional velocity model do not appear to show these features extending to depth (Fig. 2). Aftershocks and earlier seismicity²⁰ at depths below 6–7 km seem to be confined to a narrow band (see Supplementary Fig. 4). Because the fault seems straighter at depth, where large Parkfield earthquakes nucleate, than it does at the surface, it is possible that fault geometry is not the controlling factor in the location of the Parkfield segment boundaries²⁰. However, the complex surface trace geometry may result from deformation associated with the segment boundaries at seismogenic depth. That is, irregularities in the surface trace may reflect the presence of the boundaries at depth rather than being the primary cause of these boundaries.

An alternative explanation for the boundaries of the Parkfield segment is based on fault zone rheology. This segment is adjoined on the northwest by a creeping section where, perhaps, stable sliding precludes large earthquakes and on the southeast by a locked section, which may fail only in infrequent great earthquakes^{19,21}. The reasons for creep to the northwest and locking to the southeast are not clear. Properties of materials and fluid overpressure adjacent to the fault have been proposed to explain creeping and locked fault segments^{20,22}. A better knowledge of the materials and conditions within the fault zone obtained from SAFOD¹⁰ should help to discriminate between these possibilities. Ultimately, a combination of factors, including deep fault geometry, fault rheology, and stress level, may be necessary to explain why a fault creeps or is locked and what constitutes a segment boundary.

Seismic and aseismic slip

Over the long term, both seismic and aseismic slip along plate

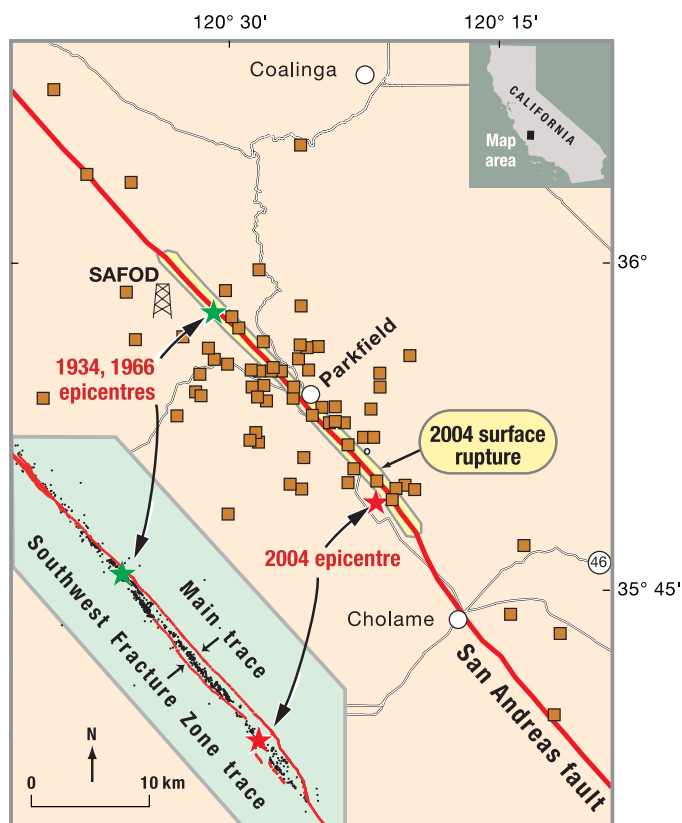


Figure 1 | Location of the 2004 Parkfield earthquake. The zone of surface rupture (yellow) is shown along the San Andreas fault (red lines). Locations of seismographs, strainmeters, creepmeters, magnetometers, and continuous GPS stations shown as squares. The strong-motion sensors (not shown here) are located on Fig. 4. Lower inset (same scale, in pale green) shows epicentres of 2004 aftershocks (black dots) plotted relative to fault traces³⁰. Upper inset, map location.

boundaries like the San Andreas fault accommodate the relative motion between the plates. To apply even the simplest mechanical model for the build-up of strain and its sudden release in earthquakes, one must account for slip on the fault that occurs aseismically. Aseismic slip has been recognized as a potentially important component of the slip budget on strike-slip faults in the San Francisco Bay area²³ and on megathrust faults along subduction plate boundaries in the Pacific Northwest region of the United States²⁴ and in Japan²⁵. How seismic and aseismic slip are distributed over a fault and how much and where aseismic slip occurs during the times between large earthquakes are, however, not well resolved. Measurements of slip beginning shortly after the 1966 Parkfield earthquake^{13,26} appear to provide some insight into these questions. The region of maximum slip in the 2004 event appears to partially fill a deficit in the distribution of slip that had accumulated beginning with the 1966 earthquake¹³, but some slip-deficient regions apparently remain (Fig. 3).

Postseismic surface slip of 35–45 cm was observed using alignment arrays following the 1966 Parkfield earthquake²⁷. The 2004 Parkfield event, however, is the first at this location for which the geodetic data were recorded during and after the earthquake with sufficient temporal and spatial resolution to enable separation of the coseismic and postseismic signals. Postseismic slip equal to about 60% of the coseismic slip occurred in the first month after the 2004 event (Fig. 3f). Alignment array data from the 2004 earthquake suggest that near-surface slip will reach 20–50 cm over the next 2–5 yr (ref. 28), comparable to what was seen after the 1966 event²⁷. Additionally, the Global Positioning System (GPS) data suggest that postseismic effects may persist for a decade, and that ultimately, the slip associated with this earthquake (coseismic plus postseismic) will balance the estimated slip deficit that existed on the fault at the time of the earthquake.

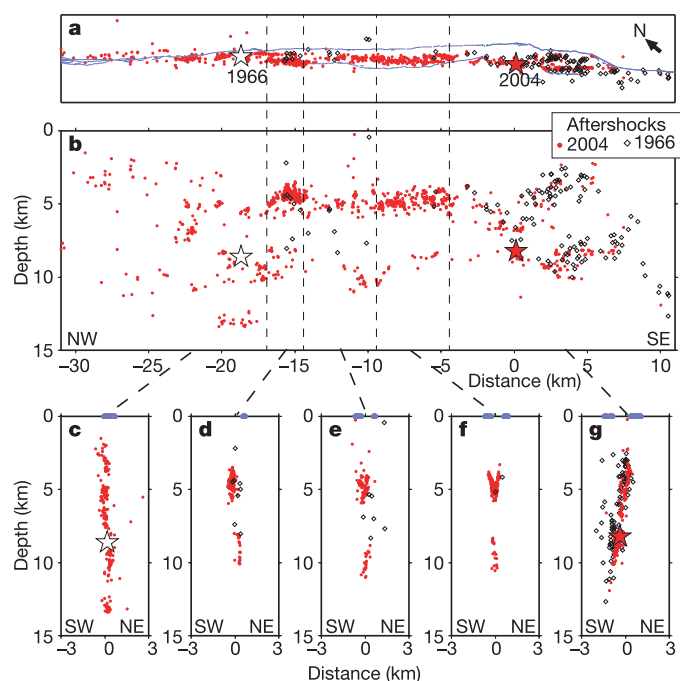


Figure 2 | Spatial distribution of Parkfield aftershocks. Locations are listed in Supplementary Table S2. Aftershocks in 2004 are shown as red dots, those in 1966 as black diamonds. Aftershocks were relocated (Supplementary Text 3) using a three-dimensional velocity model⁴⁸ and the double difference relocation technique⁴⁴. San Andreas fault traces (purple lines) and the 1966 (open star) and the 2004 (red star) main shock hypocentres are also shown. **a**, Map view. **b**, Along-fault section. **c–g**, Cross sections for the fault sections shown in **a** and **b**. The purple dots at zero depth indicate the traces shown in **a**. The aftershocks in **f** reveal multiple strands⁴⁹ activated by the main shock.

Prediction of damaging ground motion

Most of the catastrophic damage in earthquakes occurs close to the earthquake source, but relatively few recordings of strong shaking close to an earthquake have been made. The ground motion near the 2004 earthquake²⁹ was recorded at eight sites within 1 km of the rupture and at 40 sites between 1 and 10 km from the rupture (Fig. 4), nearly doubling the global data set of strong-motion records within those distances. These records show the wavefield in unprecedented detail¹⁴ and reveal large spatial variations in shaking amplitude.

The peak horizontal acceleration (PGA) for two of the records are greater than 1.0g (where g is the acceleration due to gravity) with one of these exceeding the instrument's recording capacity of 2.5g. Recordings of accelerations greater than 1g are rare, but they may not be anomalous at locations within a few kilometres of fault ruptures. Preliminary seismic slip models (see Fig. 3d for example) indicate slip was concentrated in two small regions, near the stations recording the strongest PGA. The local stress drop in these regions of concentrated slip appears to be more than an order of magnitude larger than the average stress drop of 0.2 MPa associated with the very smooth geodetic slip model (Fig. 3d). Temporal variations in rupture propagation, however, probably also influenced the radiation of the strongest shaking, as illustrated by the spatial variability in PGA (Fig. 4) and peak ground velocity close to the fault. Explaining the large variations in amplitude over distances of just a few kilometres continues to challenge our understanding of earthquake rupture dynamics and our ability to predict ground motions near the rupture (Fig. 4b and Supplementary Fig. S3).

Spatial variations in the intensity of shaking, such as peak ground acceleration and peak ground velocity, are often attributed to four factors: differences in soil conditions among sites, differences in the

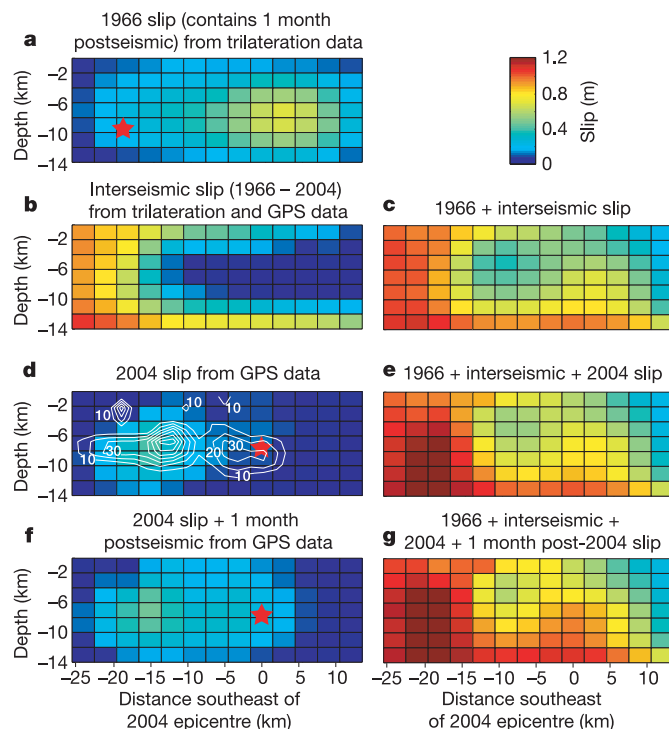


Figure 3 | Distribution of slip on the San Andreas fault since 1966 estimated from geodetic data. **a, b, d, f**, Slip; **c, e, g**, accumulated slip. Slip in the 2004 earthquake (**d**) concentrated near an area with an apparent slip deficit (compare **c** and **d**). In **d** we overlay contours of slip, estimated using both geodetic and seismologic data, giving a higher-resolution image of the slip distribution with a peak slip of 77 cm (ref. 50). These slip models illustrate how slip in earthquakes (coseismic and postseismic) combines with aseismic slip between earthquakes to generate the cumulative offset across the fault. Slip values are listed in Supplementary Table S3.

wave propagation to the sites, complexities of the rupture geometry, and heterogeneity of slip on the fault. Analysis of the strong-motion records from the 2004 earthquake should lead to a fuller understanding of how each of these factors contributes to the spatial variability in strong shaking, especially at locations close to the fault rupture. This has significant ramifications for earthquake hazard research. For example, the variability in PGA was greatest close to the rupture (Fig. 4). This suggests that complexities in the seismic source may have been the primary cause of the variations, in which case research with a greater emphasis on understanding the physical processes controlling complexity of the source would be most effective. On the other hand, near-surface soil conditions at the site

and heterogeneity in the properties of the Earth's crust that influence seismic-wave propagation are known to be important for determining the shaking at locations farther from the rupture. If these factors are found to be important for predicting the distribution of shaking within a few kilometres of the rupture as well, then directing additional resources towards developing detailed maps of these properties would also be effective. The large uncertainties in current estimates of strong ground shaking require that societal guidelines, including the Uniform Building Code and California's Alquist-Priolo Fault Zoning Act³⁰, be conservative, thereby driving up the cost of construction and hazard mitigation. To the extent that such variations in shaking are predictable, the precision of seismic hazard maps and building codes could be improved, allowing necessarily limited hazard mitigation funds to be used more effectively.

Long-term non-randomness of earthquakes

The notion that large earthquakes tend to occur as similar-size 'characteristic' events on fixed segments of a fault and that these segments are identifiable from geologic and geophysical data arose in the 1980s (refs 3, 19 and 31) and remains central to fault-based Probabilistic Seismic Hazard Analysis (PSHA). PSHA also includes other approaches such as smoothed seismicity models (see ref. 32), random events and multiple segment ruptures (see ref. 23). The sequence of earthquakes at Parkfield since 1857 has long been considered a prime example of the recurrence of a characteristic earthquake^{3,5,23}. Two classes of characteristic earthquakes have been considered for these events. In the first, events have the same faulting mechanism and magnitude, and occur on the same fault segment³¹; this class of characteristic behaviour is most appropriate for long-term forecasting of earthquakes and is often inferred from paleoseismic investigations. In the second class, the events also have the same epicentre and rupture direction³. If events in the second class were further constrained to have the same rupture time history and distribution of slip, then this class of recurrent behaviour would imply low variability in the distribution of strong ground shaking among the recurrences of characteristic events.

The 1934, 1966 and 2004 Parkfield earthquakes are remarkably similar in size (see Fig. 5 for example) and location of rupture, albeit not in epicentre or rupture propagation direction. The aftershocks of the 1966 and 2004 earthquakes delineate many of the same fault structures (Fig. 2). Furthermore, most of the observations available for the Parkfield earthquakes in 1881, 1901 and 1922 are consistent with the hypothesis that these earlier earthquakes were similar in size and general location to the later events^{3,5}. Owing to limited observations of these earlier events, a rigorous definition of the Parkfield main shocks must be limited to their overall size and their location based on rupture along the Parkfield segment³³. Michael and Jones' definition³³ was designed to encompass the Parkfield main shocks through 1966; the 2004 main shock also satisfies their definition. Thus, the Parkfield earthquakes are consistent with the first class of characteristic earthquake behaviour. However, the variability in the spatial distribution of slip for the last three events^{34,35} and the different direction of rupture propagation in the 2004 event invalidates the application of the second class of characteristic behaviour to the Parkfield earthquakes.

We note that the six Parkfield earthquakes since 1857 have occurred with statistically significant (albeit imperfect) regularity in time—more regular than random but not sufficiently periodic to be predictable in any useful way beyond long-term statistical forecasts (Supplementary Text 2). This limited regularity underlies most of the long-term prediction models proposed for the earthquakes^{3,5}. The departures from perfectly regular occurrence of these earthquakes have been interpreted using physics-based variations upon the characteristic earthquake model. For example, the Parkfield Recurrence Model, used at the outset of the experiment in 1985, assumed a constant fault loading rate and failure threshold and allowed that main shocks could be triggered early by foreshocks, but

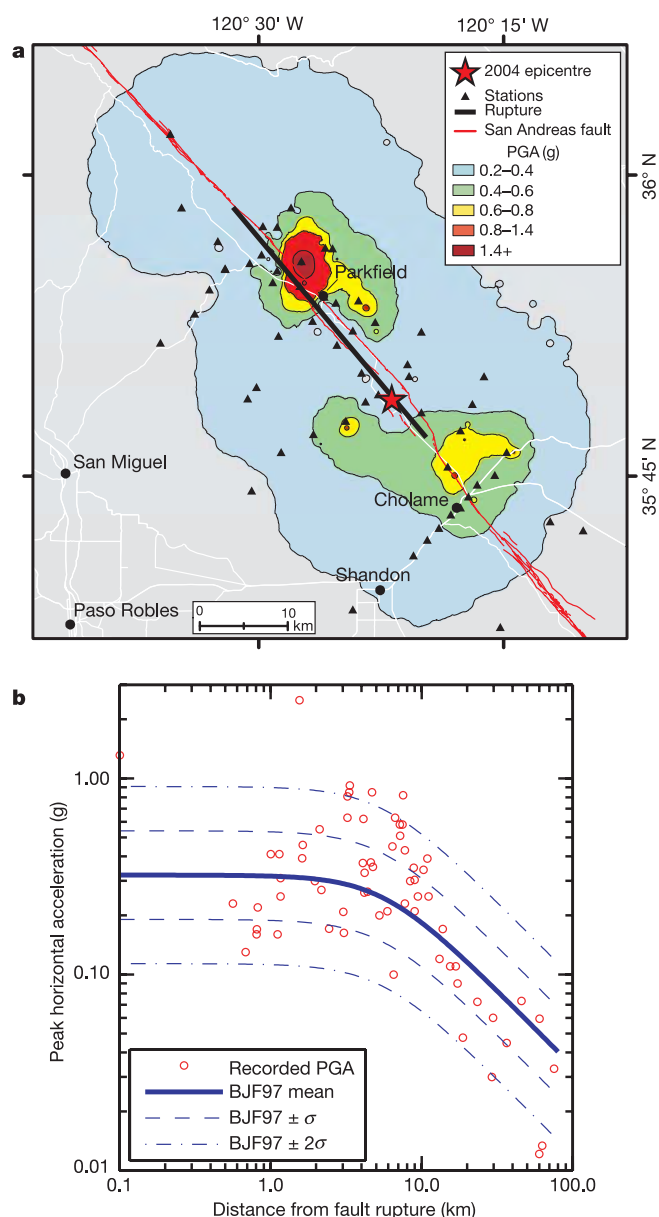


Figure 4 | Horizontal PGA from ShakeMap⁵¹. **a**, Map view. Station locations shown by triangles. The thin red line delineates the Alquist-Priolo fault traces³⁰ and the thick black line is the fault trace based on aftershock locations. **b**, Horizontal PGA as a function of distance from the fault rupture. The distances are based on the approximate projection of the fault to the ground surface. The mean (solid line) and $\pm 1\sigma$ (dashed lines) and $\pm 2\sigma$ (dash-dotted lines) for the Boore–Joyner–Fumal 1997 attenuation relation⁵² are shown. The PGA which exceeded the 2.5g limit of the instrument is plotted at 2.5g.

did not allow for late events⁵. Consequently, it has been invalidated by the long time interval between the 1966 and 2004 main shocks. In a time-predictable model, the time between successive events is proportional to the slip of the prior event; in a slip-predictable model, the size of an earthquake is proportional to the time since the prior event³⁶. Neither of these models is compatible with the sequence of Parkfield earthquake ruptures³⁴. Various forms of fault interaction have been proposed to explain the variation in recurrence intervals of the earthquakes³⁷.

Statistical models of earthquake recurrence have also been applied to these events. The variability in the time between earthquakes implies a coefficient of variability (COV) of about 0.45 which is similar to the COV used in recent forecasts for the San Francisco Bay Area²³ but greater than that proposed in earlier models for the Parkfield sequence³⁸. Earthquake activity over a wide range of smaller magnitudes ($M_w = 0$ to 5) also occurs at Parkfield. Clusters of microearthquakes that produce nearly identical waveforms repeatedly rupture small, fixed patches of the fault—some with remarkable regularity. Many of these clusters have characteristic recurrence times of months to years that scale with the magnitude of the repeating events. Changes in this recurrence time have been used to infer that slip rates over portions of the fault vary with time³⁹. Similar to the Parkfield main shocks, models of these events suggest that they may balance their local slip budget with a mix of seismic and aseismic slip⁴⁰.

The earthquakes at Parkfield, both large and small, provide a fertile

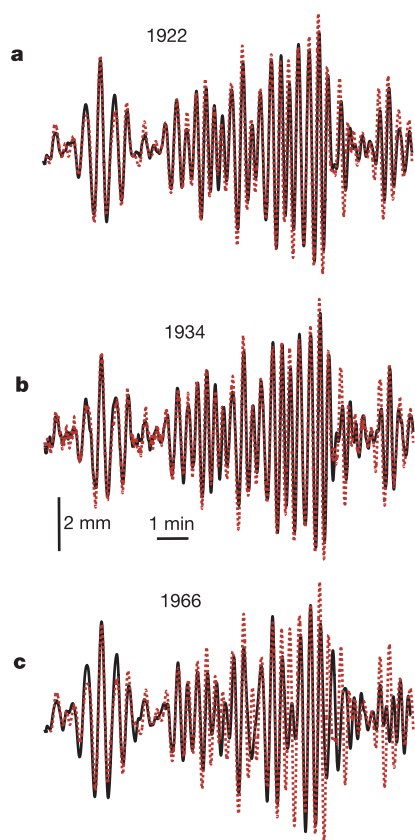


Figure 5 | Seismograms for Parkfield earthquakes at De Bilt, the Netherlands. North–south seismogram for the 2004 Parkfield earthquake (red dashed line) is plotted relative to the 1922, 1934, and 1966 Parkfield earthquakes in **a**, **b**, and **c** respectively. The 1922, 1934, and 1966 events were recorded by horizontal Galitzin seismographs. The 2004 signal, recorded by a three-component, broadband digital station located at the same site as the Galitzin seismograph, was digitally filtered to simulate a Galitzin seismograph. The similar amplitudes and waveforms imply the same seismic moment, focal mechanism, and teleseismic wave path.

laboratory for testing and refining the characteristic earthquake concept by offering information on slip distribution, rupture dynamics and afterslip, and for testing models of earthquake recurrence and interaction, which are central to contemporary earthquake hazard assessment²³. (The characteristic earthquake model can also be tested using global data sets. Kagan and Jackson⁴¹ concluded that too few of Nishenko's⁴² predicted gap-filling circum-Pacific earthquakes occurred in the first 5 yr.) Although the Parkfield earthquake history supports the use of characteristic events for earthquake forecasting, this topic remains controversial^{42,43} and we must consider whether conclusions drawn from observations at Parkfield will transfer to other seismogenic regions. For instance, does the presence of aseismic slip to the north of and within the Parkfield segment yield unusual earthquake behaviour? Observations of the large amount of postseismic slip following the 2004 earthquake suggest that it may help to balance the slip budget. Faults that do not slip aseismically may exhibit more irregular behaviour because other large events are needed to balance the slip budget. There are other faults, however, with aseismic slip that produce small repeating events and may thus produce characteristic events similar to those observed at Parkfield. These include the Hayward and Calaveras faults in California and partially coupled subduction zones^{44–46}.

Implications for future research

The magnitude and rupture extent of the 2004 Parkfield earthquake were correctly anticipated, but its time of occurrence clearly was not. This suggests that long-term earthquake forecasts require models that include higher degrees of variability (for example, see ref. 23). Although the 2004 Parkfield earthquake was ideally located within a dense monitoring network specifically designed to detect foreshocks and other possible short-term precursors, no significant signals were detected. This documented absence of clear precursory activity sets stringent bounds on the processes that preceded this earthquake. Attempts to detect short-term precursory strain changes near several other recent $M_w = 5.3$ – 7.3 earthquakes in California and Japan have also failed⁴⁷. These experiences demonstrate that reliable short-term earthquake prediction (up to a few weeks in advance) will be very difficult at best. Although the search for precursors should not be abandoned, we should thoroughly explore other ways to mitigate losses in earthquakes.

Earthquake loss mitigation begins with hazard assessment. Improved hazard assessment will require incorporating the observed variability in both earthquake sources and the resulting ground motions into probabilistic seismic hazard analysis. The history of events at Parkfield and the detailed observations of the 2004 event have revealed variability in intervals between earthquakes, variations in slip distributions of large events, and spatial variations in strong ground motion. Incorporating realistic variability into hazard assessments will entail sophisticated three-dimensional numerical models that can accurately explore many seismic cycles and include the build-up of strain via plate motions, dynamic stress changes during rupture, and postseismic deformation. Such models must be able to explain the interaction of aseismic and seismic slip, the segmentation of faults, and the strong spatial variations in the intensity of strong shaking. Complementary data from *in situ* studies of the Earth's crust, such as SAFOD¹⁰, laboratory experiments that recreate the conditions of faults in the Earth, and continued seismic monitoring will be needed to constrain the numerical models.

The value of the unique long-term record of crustal deformation being collected at Parkfield suggests that the monitoring there should continue. Parkfield-like experiments embedded within broader monitoring networks in other locations can provide similarly valuable data for faults in other tectonic contexts. Additional selected faults in California, which are already contained within sparse monitoring networks, should be densely instrumented. Large earthquakes are also anticipated on known fault segments in China, Japan, Turkey and elsewhere, and international cooperation should be

sought to develop comprehensive monitoring in these regions. The Parkfield experiment showed that diligence is required to maintain these specialized networks until a large earthquake occurs, and the detailed observations made at Parkfield demonstrate how valuable such perseverance can be for advancing our understanding of earthquakes.

Received 28 January; accepted 10 July 2005.

- Sieh, K. E. Slip along the San Andreas Fault associated with the great 1857 earthquake. *Bull. Seismol. Soc. Am.* **68**, 1421–1447 (1978).
- Toppozada, T. R., Branson, D. M., Reichle, M. S. & Hallstrom, C. L. San Andreas fault zone, California; $M \geq 5.5$ earthquake history. *Bull. Seismol. Soc. Am.* **92**, 2555–2601 (2002).
- Bakun, W. H. & McEvilly, T. V. Recurrence models and Parkfield, California, earthquakes. *J. Geophys. Res.* **89**, 3051–3058 (1984).
- Shearer, C. F. Southern San Andreas fault geometry and fault zone deformation: implications for earthquake prediction (National Earthquake Prediction Council Meeting, March, 1985). *US Geol. Surv. Open-file Rep.* **85**–507, 173–174 (USGS, Reston, Virginia, 1985).
- Bakun, W. H. & Lindh, A. G. The Parkfield, California, earthquake prediction experiment. *Science* **229**, 619–624 (1985).
- Sherburne, R. W. Ground shaking and engineering studies near the San Andreas fault zone. *Calif. Geol.* **41**, 27–32 (1988).
- Bakun, W. H. et al. Parkfield, California, earthquake prediction scenarios and response plans. *US Geol. Surv. Open-file Rep.* **87**–192, 1–45 (USGS, Reston, Virginia, 1987).
- Hill, D. P. et al. Response plans for volcanic hazards in the Long Valley caldera and Mono Craters area, California. *US Geol. Surv. Open-file Rep.* **91**–270, 1–64 (USGS, Reston, Virginia, 1991).
- USGS Advanced National Seismic System (<http://earthquake.usgs.gov/anss/>) (USGS Earthquake Hazards Program, 2000).
- Hickman, S., Zoback, M. & Ellsworth, W. Introduction to the special section: preparing for the San Andreas Fault Observatory at depth. *Geophys. Res. Lett.* **31**, L12501, doi:10.1029/2004GL020688 (2004).
- Earthscope Project *Exploring the Structure and Evolution of the North American Continent* (<http://www.earthscope.org/>) (2003).
- National Earthquake Prediction Evaluation Council Working Group. Earthquake research at Parkfield, California, 1993 and beyond—report of the NEPEC working group to evaluate the Parkfield earthquake prediction experiment. *US Geol. Surv. Circ.* **1116**, 1–14 (USGS, Reston, Virginia, 1994).
- Langbein, J. et al. Preliminary report on the 28 September 2004, $M_6.0$ Parkfield, California, earthquake. *Seismol. Res. Lett.* **76**, 10–26 (2005).
- Shakal, A. et al. Preliminary analysis of strong-motion recordings from the 28 September 2004 Parkfield, California earthquake. *Seismol. Res. Lett.* **76**, 27–39 (2005).
- Rikitake, T. *Earthquake Prediction* 7–26 (Elsevier, Amsterdam, Netherlands, 1976).
- Bakun, W. H. & McEvilly, T. V. Earthquakes near Parkfield, California; comparing the 1934 and 1966 sequences. *Science* **205**, 1375–1377 (1979).
- Nadeau, R. M. & Dolenc, D. Nonvolcanic tremors deep beneath the San Andreas fault. *Science* **307**, 389, doi:10.1126/science.1107142 (2004).
- Harris, R. A. Numerical simulations of large earthquakes: Dynamic rupture propagation on heterogeneous faults. *Pure Appl. Geophys.* **161**, 2171–2181, doi:10.1007/s00024-004-2556-8 (2004).
- Lindh, A. G. & Boore, D. M. Control of rupture by fault geometry during the 1966 Parkfield earthquake. *Bull. Seismol. Soc. Am.* **71**, 95–116 (1981).
- Eberhart-Phillips, D. & Michael, A. J. Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, Central California. *J. Geophys. Res.* **98**, 15737–15758 (1993).
- Liu, J., Klinger, Y., Sieh, K. & Rubin, C. Six similar sequential ruptures of the San Andreas Fault, Carrizo Plain, California. *Geology* **32**, 649–652 (2004).
- Irwin, W. P. & Barnes, I. Effect of geologic structure and metamorphic fluids on seismic behaviour of the San Andreas fault system in central and northern California. *Geology* **3**, 713–716 (1975).
- Working Group on California Earthquake Probabilities. Earthquake probabilities in the San Francisco Bay region: 2002–2031. *US Geol. Surv. Open-file Rep.* **03**–214, 1–340 (USGS, Reston, Virginia, 2003).
- Dragert, H., Wang, K. L. & James, T. S. A silent slip event on the deeper Cascadia subduction interface. *Science* **292**, 1525–1528 (2001).
- Kawasaki, I. et al. The 1992 Sanriku-Oki, Japan, ultra-slow earthquakes. *J. Phys. Earth* **43**, 105–116 (1995).
- Harris, R. A. & Segall, P. Detection of a locked zone at depth on the Parkfield, California segment of the San Andreas fault. *J. Geophys. Res.* **92**, 7945–7962 (1987).
- Lienkaemper, J. J. & Prescott, W. H. Historic surface slip along the San Andreas fault near Parkfield, California. *J. Geophys. Res.* **94**, 17647–17670 (1989).
- Lienkaemper, J. J., Baker, B. & McFarland, F. S. Slip in the 2004 Parkfield, California earthquake measured on alignment arrays. *Eos* **85**(47), Abstr. S54B–01 (2004).
- CISN Strong Motion Engineering Data Center (<http://www.quake.ca.gov/cisn-edc/>) (2000).
- California Geological Survey *Alquist-Priolo Earthquake Fault Zones* (<http://www.consrv.ca.gov/CGS/rghm/ap/>) (1998).
- Schwartz, D. P. & Coppersmith, K. J. Fault behaviour and characteristic earthquakes; examples from the Wasatch and San Andreas fault zones. *J. Geophys. Res.* **89**, 5681–5698 (1984).
- Frankel, A. D. et al. USGS national seismic hazard maps. *Earthq. Spectr.* **16**, 1–19 (2000).
- Michael, A. J. & Jones, L. M. Seismicity alert probabilities at Parkfield, California, revisited. *Bull. Seismol. Soc. Am.* **88**, 117–130 (1998).
- Murray, J. R. & Segall, P. Testing time-predictable earthquake recurrence by direct measure of strain accumulation and release. *Nature* **419**, 287–291 (2002).
- Segall, P. & Du, Y. How similar were the 1934 and 1966 Parkfield earthquakes? *J. Geophys. Res.* **98**, 4527–4538 (1993).
- Shimazaki, K. & Nakata, T. Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7**, 279–282 (1980).
- Ben-Zion, Y., Rice, J. R. & Dmowska, R. Interaction of the San Andreas fault creeping segment with adjacent great rupture zones, and earthquake recurrence at Parkfield. *J. Geophys. Res.* **98**, 2135–2144 (1993).
- Savage, J. C. The Parkfield Prediction Fallacy. *Bull. Seismol. Soc. Am.* **83**, 1–6 (1993).
- Nadeau, R. M. & McEvilly, T. V. Fault slip rates at depth from recurrence intervals of repeating earthquakes. *Science* **285**, 718–721 (1999).
- Beeler, N. M., Lockner, D. A. & Hickman, S. H. A simple creep-slip and stick-slip model for repeating earthquakes and its application to micro-earthquakes at Parkfield. *Bull. Seismol. Soc. Am.* **91**, 1797–1804 (2001).
- Kagan, Y. Y. & Jackson, D. D. New seismic gap hypothesis: Five years after. *J. Geophys. Res.* **100**, 3943–3959 (1995).
- Nishenko, S. P. Circum-Pacific seismic potential—1989–1999. *Pure Appl. Geophys.* **135**, 169–259 (1991).
- Main, I. *Is reliable earthquake prediction of individual earthquakes a realistic scientific goal?* (http://www.nature.com/nature/debates/earthquake/equake_contents.html) (Nature Debate 25 February to 8 April 1999).
- Waldhauser, F. & Ellsworth, W. L. A double-difference earthquake location algorithm; method and application to the northern Hayward Fault, California. *Bull. Seismol. Soc. Am.* **90**, 1353–1368 (2000).
- Schaff, D. P., Beroza, G. C. & Shaw, B. E. Post-seismic response of repeating earthquakes. *Geophys. Res. Lett.* **107**, B9, doi:10.1029/2001JB000633 (1998).
- Uchida, N., Matsuzawa, T., Igarashi, T. & Hasegawa, A. Interplate quasi-static slip off Sanriku, NE Japan, estimated from repeating earthquakes. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017452 (2003).
- Johnston, M. J. S. & Linde, A. T. Implications of crustal strain during conventional, slow, and silent earthquakes. *Int. Handbk Earthq. Eng. Seismol.* **81A**, 589–605 (2002).
- Michael, A. J. & Eberhart-Phillips, D. M. Relations among fault behaviour, subsurface geology, and three-dimensional velocity models. *Science* **253**, 651–654 (1991).
- Waldhauser, F., Ellsworth, W. L., Schaff, D. P. & Cole, A. Streaks, multiplets, and holes: high-resolution spatio-temporal behaviour of Parkfield seismicity. *Geophys. Res. Lett.* **31**, L18608, doi:10.1029/2004GL020649 (2004).
- Ji, C., Choi, K. K., King, N., Larson, K. M. & Hudnut, K. W. Co-seismic slip history and early afterslip of the Parkfield earthquake. *Eos* **85**(47), Abstr. S53D–04 (2004).
- CISN ShakeMap. *Estimated Instrumental Intensity* (<http://www.quake.ca.gov/shake/index.html>) (1997).
- Boore, D. M., Joyner, W. B. & Fumal, T. E. Equations for estimating horizontal response spectra and peak acceleration from Western North American earthquakes: A summary of recent work. *Seismol. Res. Lett.* **68**, 128–153 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The Parkfield experiment has served as a model for the collaboration of federal and state agencies with researchers in academia and industry, and many, far too numerous to list here, have contributed to its successes. In particular, J. Davis, J. Filson, A. Lindh and T. McEvilly made the experiment happen. We thank T. Hanks, S. Hough, D. Jackson, Y. Kagan, A. Lindh, M. Rymer, W. Thatcher, D. Wald and M. L. Zoback for their comments and suggestions and L. Blair, J. Boatwright, M. Huang, and D. Wald for technical assistance.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to W.H.B. (bakun@usgs.gov).

Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus

Per H. Mygind^{1*}, Rikke L. Fischer^{2*}, Kirk M. Schnorr^{1*}, Mogens T. Hansen¹, Carsten P. Sönksen¹, Svend Ludvigsen³, Dorotea Raventós¹, Steen Buskov¹, Bjarke Christensen¹, Leonardo De Maria¹, Olivier Taboureau^{1†}, Debbie Yaver⁴, Signe G. Elvig-Jørgensen¹, Marianne V. Sørensen¹, Bjørn E. Christensen¹, Søren Kjærulff¹, Niels Frimodt-Møller², Robert I. Lehrer⁵, Michael Zasloff⁶ & Hans-Henrik Kristensen¹

Animals and higher plants express endogenous peptide antibiotics called defensins. These small cysteine-rich peptides are active against bacteria, fungi and viruses. Here we describe plectasin—the first defensin to be isolated from a fungus, the saprophytic ascomycete *Pseudoplectania nigrella*. Plectasin has primary, secondary and tertiary structures that closely resemble those of defensins found in spiders, scorpions, dragonflies and mussels. Recombinant plectasin was produced at a very high, and commercially viable, yield and purity. *In vitro*, the recombinant peptide was especially active against *Streptococcus pneumoniae*, including strains resistant to conventional antibiotics. Plectasin showed extremely low toxicity in mice, and cured them of experimental peritonitis and pneumonia caused by *S. pneumoniae* as efficaciously as vancomycin and penicillin. These findings identify fungi as a novel source of antimicrobial defensins, and show the therapeutic potential of plectasin. They also suggest that the defensins of insects, molluscs and fungi arose from a common ancestral gene.

Antimicrobial peptides (AMPs) are widely distributed intrinsic host defence molecules that are probably produced by all multicellular plants and animals, as well as by some single-celled organisms^{1,2}. AMPs permit plants and animals to resist infection by environmental microbes, and represent essential components of the innate immune systems of animals³. Most AMPs are amphipathic and cationic molecules. Their net positive charge promotes their binding to the membranes of microbes, which are generally negatively charged. Many AMPs act by compromising the structural and/or functional integrity of microbial membranes^{4–6}, but some may have additional or alternative modes of action^{7–9}. The many hundreds of different AMPs described to date have been classified according to their structural features and/or their amino acid composition^{2,10}. Two families of AMPs predominate in vertebrates: cathelicidins and defensins. Cathelicidins are AMPs whose pro-peptide precursors also contain a conserved, ~100-residue-long, amino-terminal ‘cathelin’ domain. The AMP domain of different cathelicidin precursors varies widely in sequence, composition and structure. Processed cathelicidin peptides range in length from 12 to ~80 residues, and may have α -helical, β -sheet or other types of tertiary structures^{11,12}. In contrast, defensins are more uniform in their appearance. They are small cysteine-rich AMPs that mainly form β -sheet structures stabilized by three or (rarely) four conserved intramolecular cystine disulphide bridges. Three subfamilies, called α -, β - and θ -defensins, exist in vertebrates¹³. The defensin-like peptides produced by higher plants (‘plant defensins’) and by many invertebrates (‘insect defensins’) have a cystine-stabilized α -helix- β -sheet (α - β) structure

similar to the structure of many vertebrate β -defensins¹⁴. Individual defensins may show considerable selectivity, but collectively, and under optimal conditions, defensins exhibit activity against a broad range of bacteria, fungi, protozoa and enveloped viruses^{15–19}. In general terms, there are many properties that make defensins and other AMPs attractive candidates for biopharmaceutical development: they kill their target organisms rapidly; microbial resistance may occur with lower probability than observed with ‘conventional’ anti-infective agents; some exhibit adjuvant-like effects on the immune system^{20,21}; and they frequently synergize with other classes of antibiotics². However, the development of AMPs as systemic therapeutics has been hindered by several factors, including sub-optimal efficacy and safety when administered systemically, and the lack of a cost-effective means of commercial-scale production². Here we provide a detailed *in vitro* and *in vivo* characterization of plectasin, the first antimicrobial defensin peptide to be isolated from a fungus. The data show that plectasin has properties that make it an attractive candidate for development into a human therapeutic, and highlight the existence in fungi of novel peptide antibiotics with therapeutic potential.

Identification of plectasin

The fungus *P. nigrella* is a black saprophytic ascomycete found on the floor of northern European pine forests. This organism was screened for secretory proteins using a selection procedure we have applied routinely for the identification of novel bacterial and fungal enzymes²². In brief, messenger RNA was extracted from mycelia of

¹Novozymes A/S, 2880 Bagsvaerd, Denmark. ²National Center for Antimicrobials and Infection Control, Statens Serum Institut, 2300 Copenhagen S, Denmark. ³Novo Nordisk A/S, 2880 Bagsvaerd, Denmark. ⁴Novozymes Inc., Davis, California 95616, USA. ⁵Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA. ⁶Departments of Surgery and Pediatrics, Georgetown University Medical Center, Washington DC 20007, USA. [†]Present address: Novo Nordisk A/S, 2769 Maaloev, Denmark.

*These authors contributed equally to this work.

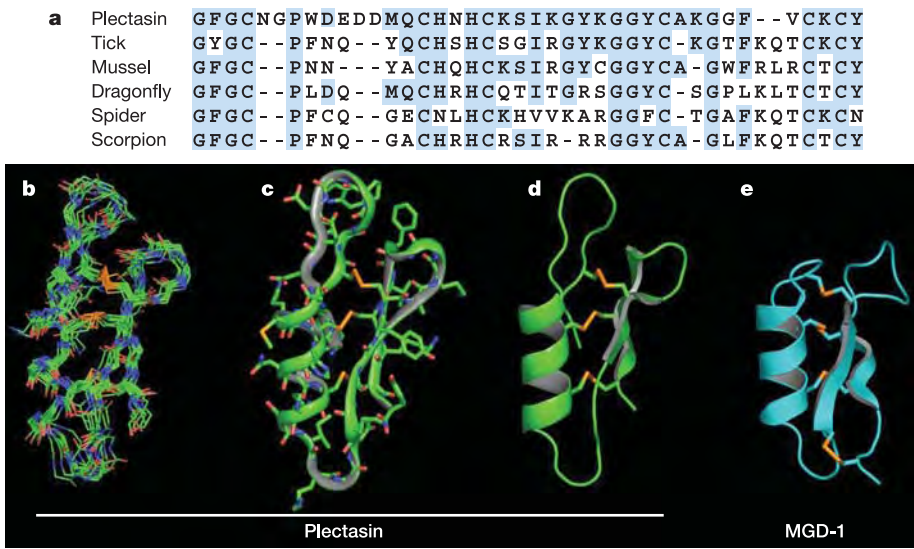


Figure 1 | Similarity of plectasin to selected invertebrate defensins.

a, Smith–Watermann local alignment of plectasin with related defensins (Tick, *Ornithodoros moubata*, Defensin A, gi:1362787; mussel, *Mytilus galloprovincialis*, MGD-1, gi:5533299; Dragonfly, *Aeschna cyanea*, Defensin, gi:259195; Spider, Defensin, gi:56579998; Scorpion, *Leiurus quinquestriatus*, Defensin, gi:399696). Some of the sequences shown are incomplete; for additional sequence data, please refer to the accession numbers provided. **b**, Ensemble of ten converged solution NMR structures of plectasin, showing the degree of accuracy of the structure. **c**, Representative structure of plectasin with emphasized backbone and with heavy atom side-chain

representation. **d**, A single representative structure of plectasin in cartoon mode showing the proximity of disulphide bridges. Cystine bridges are highlighted in orange. Structure calculation did not initially include any restraints representing a predetermined disulphide bridge pairing, however a small, but significant, improvement of convergence with one configuration of cystine bridges, that is residues 4–30, 15–37 and 19–39, gave the disulphide combination shown. **e**, Structural representation of mussel defensin MGD-1 (ref. 25; Protein Data Bank accession number 1FJN). Structural figures were produced using PyMol (ref. 46). The NMR structure of plectasin was determined as described in the Methods section.

laboratory-grown *P. nigrella*, converted into complementary DNA, and insertionally mutagenized with a transposon encoding a β -lactamase gene that lacks a signal peptide. Plasmids were introduced into *Escherichia coli*, and bacteria secreting β -lactamase selected on ampicillin plates. Plasmids were isolated from the surviving colonies

and the fungal cDNA inserts were sequenced. BLASTX and SSEARCHp sequence similarity search programs (see Supplementary Methods) against public and private databases identified a nucleotide contig containing an open reading frame encoding a peptide of 95 amino acids in length. This peptide consisted of a signal sequence

Table 1 | Antimicrobial activities of plectasin against Gram-positive bacteria

Bacterial strains*	No. tested	Minimal inhibitory concentration† (mg l ⁻¹)											
		0.063	0.125	0.25	0.5	1	2	4	8	16	32	64	≥128
<i>Streptococcus</i>													
<i>S. pneumoniae</i> PSSP	109	–	–	4	35	26	27	15	2	–	–	–	–
<i>S. pneumoniae</i> PRSP	24	–	1	4	7	5	4	3	–	–	–	–	–
<i>S. pyogenes</i> ESSP	12	2	5	3	2	–	–	–	–	–	–	–	–
<i>S. pyogenes</i> ERSP	10	1	4	4	1	–	–	–	–	–	–	–	–
Misc. (Group‡ B, C, G)	9	–	–	1	3	4	1	–	–	–	–	–	–
<i>Staphylococcus</i>													
<i>S. aureus</i> MSSA	5	–	–	–	–	–	–	1	–	3	1	–	–
<i>S. aureus</i> MRSA	5	–	–	–	–	–	–	–	–	1	3	–	1
<i>S. epidermidis</i> MSSE	5	–	–	–	–	–	–	–	3	1	1	–	–
<i>S. epidermidis</i> MRSE	5	–	–	–	–	–	–	1	2	2	–	–	–
Misc. <i>Staphylococci</i>	3	–	–	–	–	–	2	1	–	–	–	–	–
<i>Enterococcus</i>													
<i>E. faecalis</i>	6	–	–	–	–	–	–	–	–	–	–	2	4
<i>E. faecium</i> VSEF	5	–	–	–	–	–	–	–	–	1	–	4	–
<i>E. faecium</i> VREF	5	–	–	–	–	–	–	–	–	–	2	1	2
<i>Corynebacterium</i>													
<i>C. diphtheriae</i>	4	–	–	–	–	–	1	1	2	–	–	–	–
<i>C. jeikeium</i>	3	–	–	–	–	1	2	–	–	–	–	–	–
<i>Bacillus</i>													
<i>B. cereus</i>	6	–	–	–	–	–	–	–	–	–	–	–	6
<i>B. thuringiensis</i>	3	–	–	–	–	–	–	–	–	–	2	1	–
Misc. <i>Bacillus</i>	7	1	–	1	2	1	1	1	–	–	–	–	–

*Strains were from the American Type Culture Collection, the German Collection of Microorganisms and Cell Cultures, or were clinical isolates from the Statens Serum Institut.
†The minimal inhibitory concentration (MIC) was determined as described in the Methods section. *Nocardia* spp. (n = 5) and *Listeria monocytogenes* (n = 3) all showed MICs > 64 mg l⁻¹ (data not shown).
‡Lancefield antigen group.
PSSP, penicillin-sensitive *S. pneumoniae*; PRSP, penicillin-resistant *S. pneumoniae*; ESSP, erythromycin-sensitive *S. pyogenes*; ERSP, erythromycin-resistant *S. pyogenes*; MSSA, methicillin-sensitive *S. aureus*; MRSA, methicillin-resistant *S. aureus*; MSSE, methicillin-sensitive *S. epidermidis*; MRSE, methicillin-resistant *S. epidermidis*; VSEF, vancomycin-sensitive *E. faecium*; VREF, vancomycin-resistant *E. faecium*; Misc., miscellaneous.

(residues 1–23), a pro-piece (residues 24–55) and a 40-residue carboxy-terminal domain (residues 56–95) (Supplementary Fig. 1), and exhibited 50–55% sequence identity to several defensins of invertebrates (Fig. 1a). No significant sequence similarity was observed to mammalian α - or β -defensins. The putative mature peptide, plectasin, had six cysteine and five lysine residues, and a distinctive anionic tetrapeptide (DEDD) motif. Its net charge varied between +1 to +3, depending on the ionization status of its two histidines. The cDNA was transferred into an *Aspergillus oryzae* high-efficiency protein expression system²³, which yielded a primary low molecular weight secreted product (data not shown) that was further purified by cation-exchange chromatography. The measured monoisotopic mass of the product was 4,398.80 Da, as determined by liquid chromatography–electrospray ionization–quadrupole time of flight (LC–ESI–QTOF) mass spectrometry (MS) analysis. This matched the theoretical molecular weight of mature plectasin_{56–95} within 0.001%. N-terminal sequencing was consistent with the translated DNA sequence. These results provide evidence for the correct routing and processing of pre-pro-plectasin in the secretory pathway of *A. oryzae*.

Structural features of plectasin

The three-dimensional structure of plectasin was solved using proton NMR spectroscopy at pH 3.8 (Fig. 1b–d). Plectasin is poorly soluble at neutral pH ($<1 \text{ mg ml}^{-1}$), but similar structures were observed at pH 5 and pH 6 (data not shown). Structures converged to a root mean square deviation (r.m.s.d.) of 0.78 Å for backbone atoms and 1.20 Å for all heavy atoms (residues 2–39) measured relative to the geometric average structure. The first three N-terminal residues and the loop following the α -helix (residues 22–28) were poorly defined structurally, owing to difficulties in determining the geometry of several glycines from the NMR data and not necessarily due to structural disorder. Among the ten converged structures, no distance restraint violations exceeded 0.1 Å and no dihedral angle restraint violations exceeded 1°; 80.0% of the residues fell within the most favoured regions of the Ramachandran plot, 19.7% within additional allowed regions, and 0.3% within generously allowed regions. Plectasin contains an α - β motif that is comprised of an α -helix and two antiparallel β -strands, and is stabilized by three disulphide bonds (Cys4–Cys30, Cys15–Cys37 and Cys19–Cys39). The disulphide bonds determined by NMR were subsequently confirmed by a combination of MS analysis of protease digests and molecular dynamics simulations (data not shown). A similar arrangement is found in the arthropod defensins²⁴. MGD-1, a mollusc defensin that contains four disulphide bonds (Cys4–Cys25, Cys10–Cys33, Cys14–Cys35 and Cys21–Cys38), also shows structural similarity to plectasin²⁵ (Fig. 1e). From its sequence and structural features, plectasin is clearly defined as a defensin.

The similarity of plectasin to invertebrate defensins adds weight to other circumstantial evidence linking plant and insect defensins to a common ancestral gene. It is now well known that several plant and insect defensins have similar cysteine-stabilized α - β tertiary structures and manifest the same cysteine pairing motif²⁶. However, plectasin, the first fungal defensin described; shows a high degree of structural and sequence similarity to the defensins of invertebrates. It is noteworthy that the invertebrate defensins most similar to plectasin are expressed in spiders, ticks, scorpions and dragonflies—all representatives of ancient arthropod taxa^{24,27,28}. While other possibilities, such as lateral gene transfer or convergent evolution, might also account for the similarities shown in Fig. 1a, a common genetic origin for the defensins of plants, fungi and protostome invertebrates now seems highly probable. Confirmation of this hypothesis will require the characterization of additional defensin genes and peptides from other fungi and invertebrates. As the last common ancestor of plants, fungi and invertebrates is thought to have existed over one billion years ago²⁹, plectasin, while novel, is not necessarily new.

In vitro activity of plectasin

When assayed for antimicrobial activity against a broad spectrum of bacteria by standard procedures^{30,31}, plectasin showed potent activity against several species of Gram-positive bacteria. Among these were numerous clinical isolates of *S. pneumoniae*, including all 90 different serotypes and isolates resistant to clinically used antibiotics (Table 1). Gram-negative bacteria were considerably more resistant to plectasin (data not shown). Minimal inhibitory concentrations (MICs) and minimal bactericidal concentrations (MBCs) were similar in value for most isolates, suggesting that plectasin is bactericidal (data not shown). *In vitro*, *S. pneumoniae* was killed rapidly by plectasin at rates comparable to both vancomycin and penicillin (Fig. 2a). It is

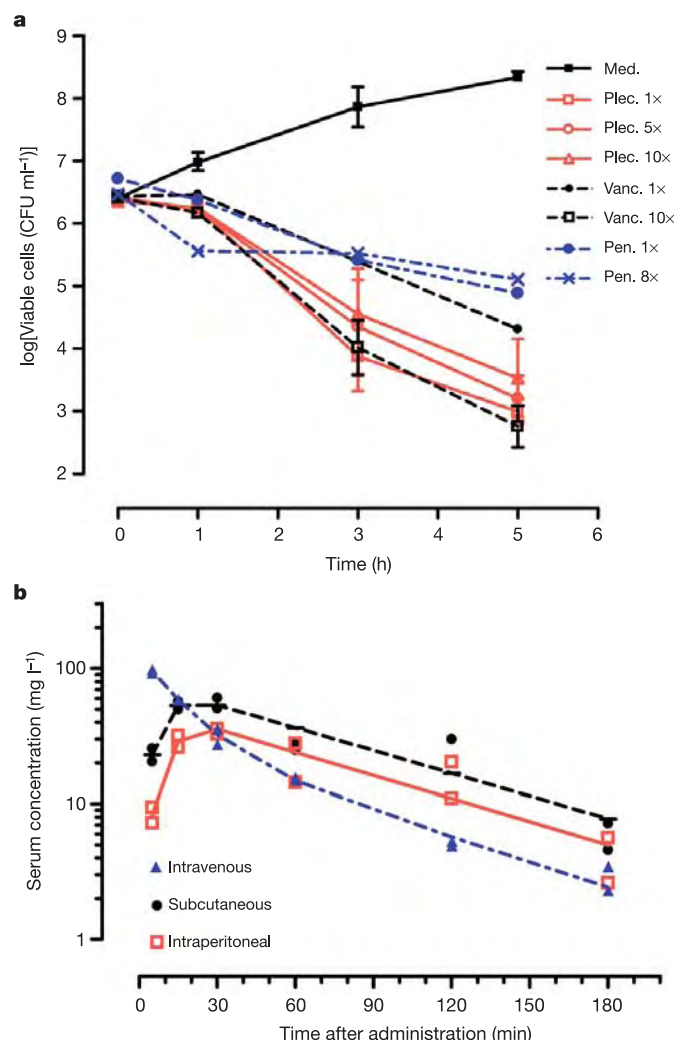


Figure 2 | In vitro and in vivo properties of plectasin. **a**, Kinetics of bacterial killing *in vitro*. *S. pneumoniae* serotypes 2 (D39), 3 (68034) and 6A (6A) were incubated in the presence of medium alone (Med.), or in the presence of either plectasin (Plec.) at 1x, 5x or 10x MIC, vancomycin (Vanc.) at 1x and 10x MIC, or penicillin (Pen.) at 1x and 8x MIC at 35 °C. Data are shown for strain D39 only. Concentrations of viable bacteria were measured as described in the Methods and Supplementary Methods sections, and the colony counts are shown as the mean and range. CFU, colony-forming units. **b**, The pharmacokinetic behaviour of plectasin *in vivo* was investigated in mice after a single dose of 14 mg per kg, using two mice at each time point. The serum concentrations were determined for intravenous, subcutaneous and intraperitoneal administration, respectively, at the time points indicated. The maximum observed mean concentrations for intravenous, subcutaneous and intraperitoneal dosing were 95 mg l⁻¹ after 5 min, and 55 mg l⁻¹ and 34 mg l⁻¹ after 30 min, respectively. The terminal half-life was estimated to 49–54 min and the concentration–time data following intravenous bolus administration showed two-compartment behaviour.

noteworthy that plectasin also is bactericidal at physiological ionic strength, especially as many vertebrate defensins require very low ionic strength *in vitro* to exert their antimicrobial effect. Plectasin was neither cytotoxic for murine L929 fibroblasts (effector concentration for half-maximum response, $EC_{50} > 512 \text{ mg l}^{-1}$) and normal human epidermal keratinocytes (NHEK) cells ($EC_{50} > 512 \text{ mg l}^{-1}$) nor haemolytic for human erythrocytes ($EC_{50} > 400 \text{ mg l}^{-1}$) (data not shown). Thus, plectasin showed considerable selectivity for bacteria over mammalian cells *in vitro*.

In vivo activity of plectasin

When given as a single intravenous injection to mice, the maximum tolerated dose of plectasin exceeded 125 mg per kg of body weight. Plectasin remained biologically active after incubation in 90% serum (either mouse or human) for 24 h at 37 °C (data not shown). After mice received a single intravenous, subcutaneous or intraperitoneal dose of plectasin (14 mg per kg), plectasin showed a two-compartment pharmacokinetic behaviour (Fig. 2b), with a terminal serum elimination half-life of 50 min and a clearance rate of about 0.1 ml min^{-1} . Intact and biologically active plectasin was recovered from mouse urine at concentrations significantly above serum values, consistent with renal clearance (data not shown).

The *in vivo* anti-infective activity of plectasin was evaluated in two mouse models of systemic pneumococcal infection. The peritonitis model³² was performed with three different pneumococcal strains

representing serotypes 2 (D39), 3 (68034) and 6A (6A). In this model, a lethal inoculum of *S. pneumoniae* was introduced into the peritoneum. The effect of treatment was evaluated in two different ways: (1) a reduction in peritoneal bacterial counts in actively treated as compared with untreated control mice, or (2) survival for up to seven days after active treatment as compared with untreated controls. Figure 3a shows the results of a study where inoculated mice were treated 1 h later with either plectasin (10 mg per kg, intravenously) or vancomycin (70 mg per kg, subcutaneously), whereas the control group was left untreated. Peritoneal bacterial counts were performed at 0 h, 2 h and 5 h after treatment (Fig. 3a). Administration of plectasin caused the concentration of viable pneumococci in the peritoneum to fall 10-fold and 1,000-fold at 2 h and 5 h, respectively. This reduction is paralleled by a similar reduction in the blood (data not shown). In the untreated controls, the concentration of viable pneumococci in the peritoneum increased >10-fold by 5 h. The survival of plectasin-treated mice was investigated in this model and compared with control groups of vehicle-treated mice. The mice were inoculated with the D39 or 68034 strains (time, -1 h), and 1 h later (time, 0 h) treated subcutaneously with 10 mg per kg of plectasin in groups of eight mice. The following dosing regimens were investigated: plectasin at 0 h (only for strain D39), 0 h and 5 h, or 0 h, 5 h, 24 h and 29 h. The mice were inspected for seven days, and if severe pain and disease was observed during the study period the mice were killed and scored as dead by the infection. At the end of the

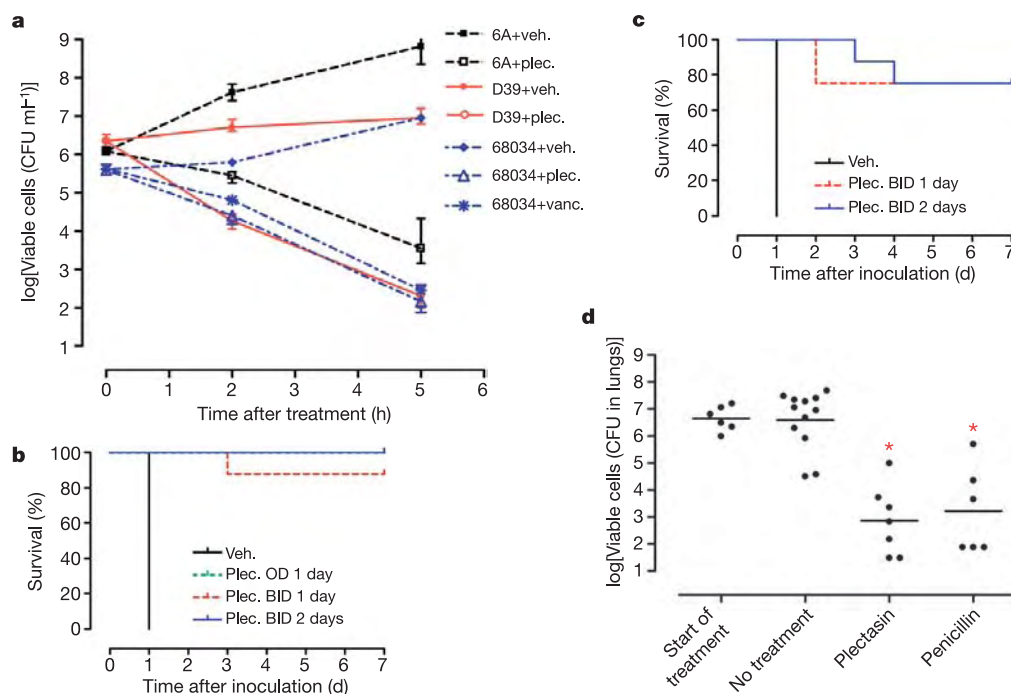


Figure 3 | In vivo properties of plectasin. **a**, Treatment of pneumococcal peritoneal infection. Three mice per sampling point were inoculated intraperitoneally as described in the Supplementary Methods section, and were treated 1 h later with a single dose of 10 mg per kg plectasin or 70 mg per kg vancomycin (vancomycin data shown only for strain 68034). Peritoneal wash was sampled for microbial counts before treatment, and at 2 h and 5 h after treatment. Colony counts are shown as the mean and range. The colony counts in mice treated with plectasin or vancomycin were significantly lower ($P < 0.0001$) compared to the group of vehicle-treated (veh.) mice. **b**, Survival after peritoneal infection with *S. pneumoniae* D39. Treatment with plectasin subcutaneously at 10 mg per kg was initiated 1 h after inoculation. The following dosing regimens were investigated in groups of eight mice: vehicle treated, plectasin once daily (OD) for one day, twice daily (BID) for one day, or BID for two days. All mice treated with plectasin survived the seven-day study period, except for one mouse in the group receiving the BID for one day treatment. The survival fraction of the vehicle-treated control group was 0 out of 8 mice ($P < 0.0002$). **c**, Survival after

peritoneal infection with *S. pneumoniae* 68034. Treatment with plectasin subcutaneously at 10 mg per kg was initiated 1 h after inoculation. The following dosing regimens were investigated in groups of eight mice: vehicle treatment, or plectasin BID for one or two days. The survival fraction at seven days post-infection was 6 out of 8 for plectasin-treated mice and 0 out of 8 for vehicle-treated mice ($P < 0.0002$). **d**, Treatment of pneumococcal pneumonia. Six to seven mice per sampling point were inoculated with *S. pneumoniae* (data shown only for strain D39) through the nasopharynx, as described in the Supplementary Methods section, and 24 h later the animals received plectasin at 10 mg per kg or penicillin at 30 mg per kg. One day after treatment the animals were necropsied, the lungs were removed, and the level of *S. pneumoniae* present in lung homogenates were determined, as described in the Supplementary Methods section. Each point represents the determination from a single animal and the line shows the mean log value. An asterisk denotes significant reduction in colony counts compared to the start of treatment ($P < 0.003$).

study, all surviving mice were unaffected and behaved normally. Treatment with plectasin cured all mice infected with strain D39 strain, except for one mouse treated twice for one day (Fig. 3b). The survival fraction of mice infected with strain 68034 was 6 out of 8 for both dosing regimens tested (Fig. 3c).

The second model was a mouse pneumonia model³³, and used the same *S. pneumoniae* strains tested in the first model (Fig. 3d). Animals received an inoculum of *S. pneumoniae* intranasally, and were left untreated for the next 24 h in order for pneumonia to develop. The next day, the treated animals received a single intravenous dose of plectasin (10 mg per kg) or two doses of penicillin (totalling 30 mg per kg). All animals were euthanized the following day. The lungs were removed, and viable pneumococci within the lung tissue were counted. Plectasin or penicillin treatment resulted in median bacterial counts at least 1,000-fold to 10,000-fold lower than those observed in untreated controls. Overall, the efficacy of plectasin was comparable to vancomycin in the peritoneal model (Fig. 3a) and to penicillin in the pneumonia model (Fig. 3d).

Discussion

Plectasin kills *S. pneumoniae* *in vitro* at rates comparable to penicillin and vancomycin (Fig. 2a), antibiotics that perturb proteoglycan biosynthesis^{34,35}. The concentration of plectasin that stops bacterial growth *in vitro* (the MIC) is about the same as that required to kill the organisms (the MBC). This property suggests that once the peptide has altered the microbe's growth *in vitro*, an irreversible process leading to death begins to occur. Many of the antimicrobial peptides from multicellular organisms act by binding to cellular membranes and directly perturb membrane function³⁶. Such antimicrobial peptides typically kill target microbes within seconds to minutes of exposure, at a concentration equivalent to the MIC (ref. 2). The slower kinetics of killing by plectasin suggests that alternative mechanisms should be considered to explain its mode of action. The precise mechanism by which plectasin exerts its antimicrobial activity is currently under investigation.

The discovery of plectasin has implications for the development of defensins as anti-infective therapeutics. Commercial development of defensins has been hindered, in large part, by the technical challenge of producing them at the scale and purity required for a pharmaceutical product. However, fully processed active plectasin can effectively be produced at high yields in a proprietary fungal expression system that is currently used for industrial scale, cost effective production of other proteins^{37,38}.

S. pneumoniae is a leading cause of meningitis, community-acquired pneumonia, sepsis and otitis media³⁹. Increasing bacterial resistance to conventional antibiotics threatens the future of many antibiotics in current use⁴⁰. In our initial studies, plectasin did not exhibit cross-resistance with other conventionally used antibiotics; for example, penicillin, cefotaxime, erythromycin, tetracycline, chloramphenicol, clindamycin or trimethoprim-sulphamethoxazole. Historically, the 'Antibiotic Era' was born almost 80 years ago when an uninvited *Penicillium* spore landed on a Petri dish and eliminated nearby staphylococcal colonies⁴¹. Perhaps its duration can be extended by antimicrobial peptides, such as plectasin, produced by other fungi.

METHODS

Construction of the *P. nigrella* cDNA library. *P. nigrella* mycelia, cultured for 7 d at 28 °C on potato dextrose agar, were inoculated into Mex1 medium (20 g l⁻¹ soybean, 15 g l⁻¹ wheat bran, 10 g l⁻¹ cellulose Avicel, 5 g l⁻¹ maltodextrin, 3 g l⁻¹ bactopeptone, 0.2 g l⁻¹ Pluronic PE6100 and 1 g l⁻¹ olive oil). After growing for 5 d at 28 °C at 250 r.p.m., mycelia were recovered, frozen in liquid nitrogen, and stored at -80 °C until used. Total RNA was extracted from powdered *P. nigrella* mycelia as previously described⁴². Poly-A⁺-enriched mRNA was recovered by oligo(dT)-cellulose affinity chromatography. Double-stranded cDNA was generated using the purified mRNA and a polyA⁺-*NotI* primer. After treatment with mung bean nuclease, the cDNA was blunt ended with T4 DNA polymerase, ligated to *EcoRI* adaptors (Invitrogen), cleaved with the restriction

endonuclease *NotI*, and purified using the GFX DNA isolation kit (GE Healthcare). The cDNA pool was directionally cloned into an *EcoRI/NotI*-digested plasmid. This plasmid contained *EcoRI/NotI* restriction sites proximal to the Shine-Dalgarno region of the *lac* promoter, allowing *EcoRI/NotI*-adapted cDNA sequences to be cloned into the vector and the resulting constructs to be transcribed and translated in an *E. coli* host. The ligation mixture was electroporated into *E. coli* DH10B cells (Invitrogen) and plated onto Luria-Bertani agar plates with 50 mg l⁻¹ kanamycin. In total, 30,000 *E. coli* colonies were scraped off the selection plates and plasmid DNA prepared with ion exchange column chromatography (Qiagen). Transposon-assisted signal trapping was performed on the resulting cDNA plasmid library as described in the Supplementary Information and in ref. 22.

Cloning and recombinant protein expression. The plectasin coding sequence was amplified from cDNA using the Expand high fidelity polymerase chain reaction (PCR) system (Roche) with plectasin-specific oligonucleotides (primer 1: 5'-TCTGGATCCACCATGCAATTTACCACCATCTCTC-3'; primer 2: 5'-TCTCTCGAGCTAGTAACACTTGCAACAAAGC-3'). The 309-base-pair (bp) PCR product was digested with *Bam*HI and *Xho*I, which cut in the overhangs introduced by the PCR primers. The digested DNA fragment was cloned into an *Aspergillus* expression plasmid (pMT2188) based on pCaHj527 (ref. 23). The resulting plasmid (pMT2548) contained a full-length plectasin cDNA gene, and was transformed into the *A. oryzae* strain BECh2. Thirty transformants were re-isolated twice under selective and non-inducing conditions on Cove minimal plates with sucrose and acetamide. To test the expression and secretion of plectasin, transformants were grown for 6 d at 30 °C in YPM media (2% (w/v) peptone, 1% (w/v) yeast extract, 2% (w/v) maltose). Supernatants were run on NuPage 10% Bis-Tris SDS gels (Invitrogen) with MES running buffer to allow separation in the low molecular weight range.

Purification of recombinant plectasin. The post-fermentation broth was sterile-filtered, adjusted to pH 6, diluted to a conductivity of 7 mS cm⁻¹, applied to a carboxymethyl-Sephadex column (GE Healthcare), and eluted with a 0–1 M NaCl gradient in 50 mM malonic acid. Fractions containing plectasin were desalted and analysed by matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) MS. Analytical-scale reverse-phase-high-performance liquid chromatography (RP-HPLC) was performed on a Jupiter C18 (5 µm, 300 Å, 150 mm × 2 mm) column (Phenomenex). The column was eluted over 50 min with an 8 to 80% gradient of acetonitrile in 0.1% TFA at 0.15 ml min⁻¹. N-terminal sequencing was done on a Procise automatic sequencer (Applied Biosystems) and the concentration was determined by amino-acid analysis.

Molecular and structural characterization. MALDI-TOF MS analysis was performed by the dried-droplet method on a Voyager DE Pro mass spectrometer (Applied Biosystems), run in linear mode with positive ionization, with external calibration using CalMix2 (Applied Biosystems). LC-ESI-QTOF analysis was performed on a QTOF mass analyser (Waters-Micromass) coupled to an Ultimate HPLC system (LC-Packings). NMR spectra of purified plectasin (1–2.5 mM in 30 mM acetic acid, pH 3.8) were recorded at 600 MHz using Varian Inova at 27 °C as previously described⁴³.

Structure calculation protocol. The structure calculation program used was CNX version 2000.1 (Molecular Simulations, Yale University) licensed from Accelrys. The input file used was 'dg_sa.inp', essentially using a standard distance geometry and simulated annealing protocol. NMR spectra were processed using Xwinnmr (Bruker Spectrospin) and NMR spectral assignments and cross-peak book-keeping was performed using PRONTO (ref. 44). Structures were analysed visually and geometrically using CNX and Procheck (ref. 45).

Antimicrobial testing. MICs were determined by the microdilution broth method according to National Committee for Clinical Laboratory Standards/Clinical and Laboratory Standards Institute (NCCLS/CLSI) guidelines³⁰. In brief, fresh overnight colonies were suspended to a turbidity of 0.5 McFarland units and further diluted in Mueller-Hinton BBLII broth (Becton Dickinson). The broth was supplemented with 2–5% defibrinated horse blood for all *Streptococcus*, *Listeria*, *Neisseria*, *Bacillus* and *Corynebacterium* species. The diluted bacterial suspensions were added to wells containing twofold peptide dilutions. The polypropylene trays (Nunc) were incubated at 35 °C in ambient air for 16–20 h for *Staphylococcus*, *Enterococcus* and *Listeria* spp., 24 h for *Bacillus*, *Corynebacterium* and *Neisseria* spp. (in 5% CO₂) and *Streptococcus* spp., and 48 h for *Nocardia* spp.

Received 24 February; accepted 20 July 2005.

1. Tossi, A., Sandri, L. & Giangaspero, A. Amphipathic, α -helical antimicrobial peptides. *Biopolymers* 55, 4–30 (2000).
2. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* 415, 389–395 (2002).
3. Boman, H. G. Innate immunity and the normal microflora. *Immunol. Rev.* 173, 5–16 (2000).

4. Matsuzaki, K. Why and how are peptide–lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes. *Biochim. Biophys. Acta Biomembr.* **1462**, 1–10 (1999).
5. Yang, L., Weiss, T. M., Lehrer, R. I. & Huang, H. W. Crystallization of antimicrobial pores in membranes: magainin and protegrin. *Biophys. J.* **79**, 2002–2009 (2000).
6. Shai, Y. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α -helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta Biomembr.* **1462**, 55–70 (1999).
7. Park, C. B., Kim, H. S. & Kim, S. C. Mechanism of action of the antimicrobial peptide buforin II: buforin II kills microorganisms by penetrating the cell membrane and inhibiting cellular functions. *Biochem. Biophys. Res. Commun.* **244**, 253–257 (1998).
8. Kragol, G. *et al.* The antibacterial peptide pyrrocorin inhibits the ATPase actions of DnaK and prevents chaperone-assisted protein folding. *Biochemistry* **40**, 3016–3026 (2001).
9. Del Castillo, F. J., Del Castillo, I. & Moreno, F. Construction and characterization of mutations at codon 751 of the *Escherichia coli* *gyrB* gene that confer resistance to the antimicrobial peptide microcin B17 and alter the activity of DNA gyrase. *J. Bacteriol.* **183**, 2137–2140 (2001).
10. Gennaro, R., Zanetti, M., Benincasa, M., Podda, E. & Miani, M. Pro-rich antimicrobial peptides from animals: structure, biological functions and mechanism of action. *Curr. Pharm. Des.* **8**, 763–778 (2002).
11. Lehrer, R. I. & Ganz, T. Cathelicidins: a family of endogenous antimicrobial peptides. *Curr. Opin. Hematol.* **9**, 18–22 (2002).
12. Zanetti, M. Cathelicidins, multifunctional peptides of the innate immunity. *J. Leukoc. Biol.* **75**, 39–48 (2004).
13. Lehrer, R. I. Primate defensins. *Nature Rev. Microbiol.* **2**, 727–738 (2004).
14. Ganz, T. Defensins: antimicrobial peptides of innate immunity. *Nature Rev. Immunol.* **3**, 710–720 (2003).
15. Selsted, M. E., Szklarek, D. & Lehrer, R. I. Purification and antibacterial activity of antimicrobial peptides of rabbit granulocytes. *Infect. Immun.* **45**, 150–154 (1984).
16. Selsted, M. E., Szklarek, D., Ganz, T. & Lehrer, R. I. Activity of rabbit leukocyte peptides against *Candida albicans*. *Infect. Immun.* **49**, 202–206 (1985).
17. Lehrer, R. I., Daher, K., Ganz, T. & Selsted, M. E. Direct inactivation of viruses by MCP-1 and MCP-2, natural peptide antibiotics from rabbit leukocytes. *J. Virol.* **54**, 467–472 (1985).
18. Daher, K., Selsted, M. E. & Lehrer, R. I. Direct inactivation of viruses by human granulocyte defensins. *J. Virol.* **60**, 1068–1074 (1986).
19. Cole, A. M. Minidefensins and other antimicrobial peptides: candidate anti-HIV microbicides. *Expert Opin. Ther. Targets* **7**, 329–341 (2003).
20. Yang, D. *et al.* β -defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science* **286**, 525–528 (1999).
21. Biragyn, A. *et al.* Toll-like receptor 4-dependent activation of dendritic cells by β -defensin 2. *Science* **298**, 1025–1029 (2002).
22. Becker, F. *et al.* Development of *in vitro* transposon assisted signal sequence trapping and its use in screening *Bacillus halodurans* C125 and *Sulfolobus solfataricus* P2 gene libraries. *J. Microbiol. Methods* **57**, 123–133 (2004).
23. Schnorr, K., Hansen, M. T., Mygind, P. H., Segura, D. R. & Kristensen, H.-H. Novel plectasin polypeptide having antimicrobial activity, useful for (for example) killing or inhibiting microbial cell growth, for use as a medicament and as antimicrobial therapeutic or prophylactic agent. International patent application WO2003044049–A1 (2003).
24. Froy, O. & Gurevitz, M. Arthropod defensins illuminate the divergence of scorpion neurotoxins. *J. Pept. Sci.* **10**, 714–718 (2004).
25. Yang, Y. S. *et al.* Solution structure and activity of the synthetic four-disulfide bond Mediterranean mussel defensin (MGD-1). *Biochemistry* **39**, 14436–14447 (2000).
26. Landon, C., Sodano, P., Hetru, C., Hoffmann, J. & Ptak, M. Solution structure of drosomycin, the first inducible antifungal protein from insects. *Protein Sci.* **6**, 1878–1884 (1997).
27. Volkoff, A. N. *et al.* Characterization and transcriptional profiles of three *Spodoptera frugiperda* genes encoding cysteine-rich peptides. A new class of defensin-like genes from lepidopteran insects? *Gene* **319**, 43–53 (2003).
28. Bulet, P. *et al.* A novel insect defensin mediates the inducible antibacterial activity in larvae of the dragonfly *Aeschna cyanea* (Paleoptera, Odonata). *Eur. J. Biochem.* **209**, 977–984 (1992).
29. Doolittle, R. F., Feng, D. F., Tsang, S., Cho, G. & Little, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **271**, 470–477 (1996).
30. The Clinical and Laboratory Standards Institute (formerly the National Committee for Clinical Laboratory Standards). Guideline M7–A6: Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. (<http://www.nccls.org>).
31. Steinberg, D. A. & Lehrer, R. I. in *Methods in Molecular Biology* Vol. 78 (ed. Schafer, W. M.) 169–186 (Humana, Totowa, New Jersey, 1997).
32. Frimodt-Møller, N., Knudsen, J. D. & Espersen, F. in *Handbook of Animal Models of Infection* (eds Zak, O. & Sande, M. A.) 127–136 (Academic, London, 1999).
33. Erlendsdottir, H. *et al.* Penicillin pharmacodynamics in four experimental pneumococcal infection models. *Antimicrob. Agents Chemother.* **45**, 1078–1085 (2001).
34. Bayles, K. W. The bactericidal action of penicillin: New clues to an unsolved mystery. *Trends Microbiol.* **8**, 274–278 (2000).
35. Williams, D. H. & Bardsley, B. The vancomycin group of antibiotics and the fight against resistant bacteria. *Angew. Chem. Int. Edn Engl.* **38**, 1173–1193 (1999).
36. Nakajima, Y. *et al.* Antibacterial activity and mechanism of action of tick defensin against Gram-positive bacteria. *Biochim. Biophys. Acta* **1624**, 125–130 (2003).
37. Hjort, C. M. in *Genetically Engineered Food: Methods and Detection* (ed. Heller, K. J.) 86–99 (Wiley, 2003).
38. Yoder, W. T. & Lehmbeck, J. in *Advances in Fungal Biotechnology for Industry, Agriculture, and Medicine* (eds Tkacz, J. S. & Lange, L.) 201–219 (Kluwer Academic/Plenum, New York, 2004).
39. The Advisory Committee on Immunization Practices. Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). *Morb. Mortal. Wkly Rep.* **46**, 1–24 (1997).
40. Klugman, K. P. Bacteriological evidence of antibiotic failure in pneumococcal lower respiratory tract infections. *Eur. Respir. J.* **20** (suppl.), 3s–8s (2002).
41. Fleming, A. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenza*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
42. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299 (1979).
43. Ludvigsen, S., Thim, L., Blom, A. M. & Wulff, B. S. Solution structure of the satiety factor, CART, reveals new functionality of a well-known fold. *Biochemistry* **40**, 9082–9088 (2001).
44. Kjær, M., Andersen, K. V. & Poulsen, F. M. Automated and semiautomated analysis of homo- and heteronuclear multidimensional nuclear magnetic resonance spectra of proteins: the program PRONTO. *Methods Enzymol.* **239**, 288–307 (1994).
45. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
46. DeLano, W. L. *The PyMOL Molecular Graphics System* (<http://www.pymol.org>) (DeLano Scientific, San Carlos, California, 2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. Ellingsgaard, A. L. Hansen, M. Markvardsen, J. Theil, A. Blom, B. Cherry, S. Otani and F. Hansen for expert technical assistance.

Author Information Plectasin DNA sequence has been submitted to EMBL with the accession numbers EMBL:AJ964941 and TREMBL:CAI83768. Plectasin coordinates have been deposited to the Protein Data Bank under accession number 1ZFU, and relevant NMR data will be submitted to the Biological Magnetic Resonance Data Bank. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.-H.K. (hahk@novozymes.com).

The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators

Rong-Gui Hu^{1*}, Jun Sheng^{1*}, Xin Qi², Zhenming Xu^{1†}, Terry T. Takahashi² & Alexander Varshavsky¹

The conjugation of arginine to proteins is a part of the N-end rule pathway of protein degradation. Three amino (N)-terminal residues—aspartate, glutamate and cysteine—are arginylated by *ATE1*-encoded arginyl-transferases. Here we report that oxidation of N-terminal cysteine is essential for its arginylation. The *in vivo* oxidation of N-terminal cysteine, before its arginylation, is shown to require nitric oxide. We reconstituted this process *in vitro* as well. The levels of regulatory proteins bearing N-terminal cysteine, such as RGS4, RGS5 and RGS16, are greatly increased in mouse *ATE1*^{−/−} embryos, which lack arginylation. Stabilization of these proteins, the first physiological substrates of mammalian N-end rule pathway, may underlie cardiovascular defects in *ATE1*^{−/−} embryos. Our findings identify the N-end rule pathway as a new nitric oxide sensor that functions through its ability to destroy specific regulatory proteins bearing N-terminal cysteine, at rates controlled by nitric oxide and apparently by oxygen as well.

The N-end rule relates the *in vivo* half-life of a protein to the identity of its N-terminal residue^{1–4}. A ubiquitin-dependent pathway, called the N-end rule pathway, recognizes degradation signals (degrons) that include the signals called N-degrons (Fig. 1a). An N-degron consists of a protein's destabilizing N-terminal residue and an internal Lys residue. The latter is the site of formation of a poly-ubiquitin chain^{5–8}. The N-end rule has a hierarchic structure. N-terminal Asn and Gln are tertiary destabilizing residues in that they function through their deamidation, by N-terminal amidohydrolases^{9,10}, to yield the secondary destabilizing residues Asp and Glu. The activity of N-terminal Asp and Glu requires their conjugation, by *ATE1*-encoded isoforms of Arg-tRNA-protein transferase (R-transferase), to Arg, one of the primary destabilizing residues^{11,12}. The latter are recognized by E3 ubiquitin ligases of the N-end rule pathway^{2,4,13,14}. In mammals, destabilizing N-terminal residues that function through their arginylation are not only Asp and Glu but also Cys (Fig. 1a), which is a stabilizing (unarginylated) residue in the yeast *Saccharomyces cerevisiae*^{11,12,15}. Known functions of the N-end rule pathway include the control of peptide import (through conditional degradation of the import's repressor)^{13,16}, the regulation of apoptosis (through degradation of a caspase-processed inhibitor of apoptosis)^{3,17} and the fidelity of chromosome segregation (through degradation of a conditionally produced cohesin's fragment)¹⁸, as well as the regulation of meiosis, cardiovascular development in animals and leaf senescence in plants (refs 4, 12 and references therein).

Nitric oxide (NO) is produced in eukaryotes largely by NO synthases. This compound and its derivatives have a function, as either stressors or regulators, in a vast range of processes, including cardiovascular homeostasis, immunity, neurotransmission, ion conductance, glycolysis and apoptosis (reviewed in refs 19–26). Biological effects of NO are mediated by its covalent modifications of proteins, either of their prosthetic groups or amino-acid residues,

particularly Cys and Tyr. The reactivity of these residues towards NO is modulated by their sequence contexts^{20–23,27–29}. NO converts Cys residues to S-nitrosothiols, a process that can involve oxygen or its derivatives. S-Nitrosylation modulates protein functions either directly or after additional (often oxygen-dependent) transformations that yield oxidized Cys, such as Cys-sulphinic acid (CysO₂H) or Cys-sulphonic acid (CysO₃H) (refs 20–22).

An N-degron is produced from pre-N-degron through a proteolytic cleavage. Methionine aminopeptidases (MetAPs) remove Met from the N terminus of a newly formed protein if, and only if, the residue at position 2, to be made N-terminal after cleavage, has a small enough side chain². Consequently, of the 13 destabilizing residues of the mammalian N-end rule (Fig. 1a), only Cys can be made N-terminal by MetAPs (Fig. 1b). (Any destabilizing residue, including Cys, can be made N-terminal through internal cleavages of proteins by other proteases, such as separases, caspases and calpains^{3,17,18} (Fig. 1a).) Our previous work showed that Cys at position 2 of the RGS4 protein arginylated *in vivo* was CysO₃H, rather than Cys, suggesting that the oxidation of N-terminal Cys might precede arginylation, and might be required for it¹².

We now show that oxidation of N-terminal Cys is essential for its arginylation. We also discovered that the oxidation of a protein's N-terminal Cys *in vivo*, before its arginylation, requires NO. This explains why N-terminal Cys is a destabilizing residue in mammalian cells^{2,12,15}, which produce NO, but stabilizing in yeast⁵, which lack NO synthases. We reconstituted the NO-dependent arginylation of N-terminal Cys in an *in vitro* system as well. This process requires a basic residue at position 2 of a substrate. The levels of regulatory proteins with this N-terminal motif (Cys-(basic residue)), such as RGS4, RGS5 and RGS16, are greatly increased in mouse *ATE1*^{−/−} embryos, which lack arginylation. RGS4, RGS5 and RGS16 are the first physiological substrates of the mammalian N-end rule pathway. Given the involvement of these proteins in cardiovascular

¹Division of Biology, and ²Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA. †Present address: Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697-4120, USA.

*These authors contributed equally to this work.

homeostasis^{30,31} and tubulogenesis³², their stabilization might underlie the previously observed abnormal angiogenesis and heart defects in *ATE1*^{-/-} embryos¹². A mammalian genome encodes approx. 30 proteins, including RGS4, RGS5 and RGS16, that contain the Met-Cys-(basic residue) N-terminal motif, which acts as a MetAP-cleaved, NO-dependent, arginylation-mediated, Cys-containing pre-N-degron. Taken together, our results identify the arginylation branch of the N-end rule pathway as a new sensor of NO in

mammalian cells that functions through its ability to destroy specific regulatory proteins bearing N-terminal Cys, at rates controlled by NO and apparently by oxygen as well.

Oxidation of N-terminal cysteine is required for its arginylation

We wished to determine whether the presence of CysO₃H (instead of Cys) at position 2 of the RGS4 protein arginylated *in vivo*¹² reflected the requirement for oxidation of Cys before its arginylation, as

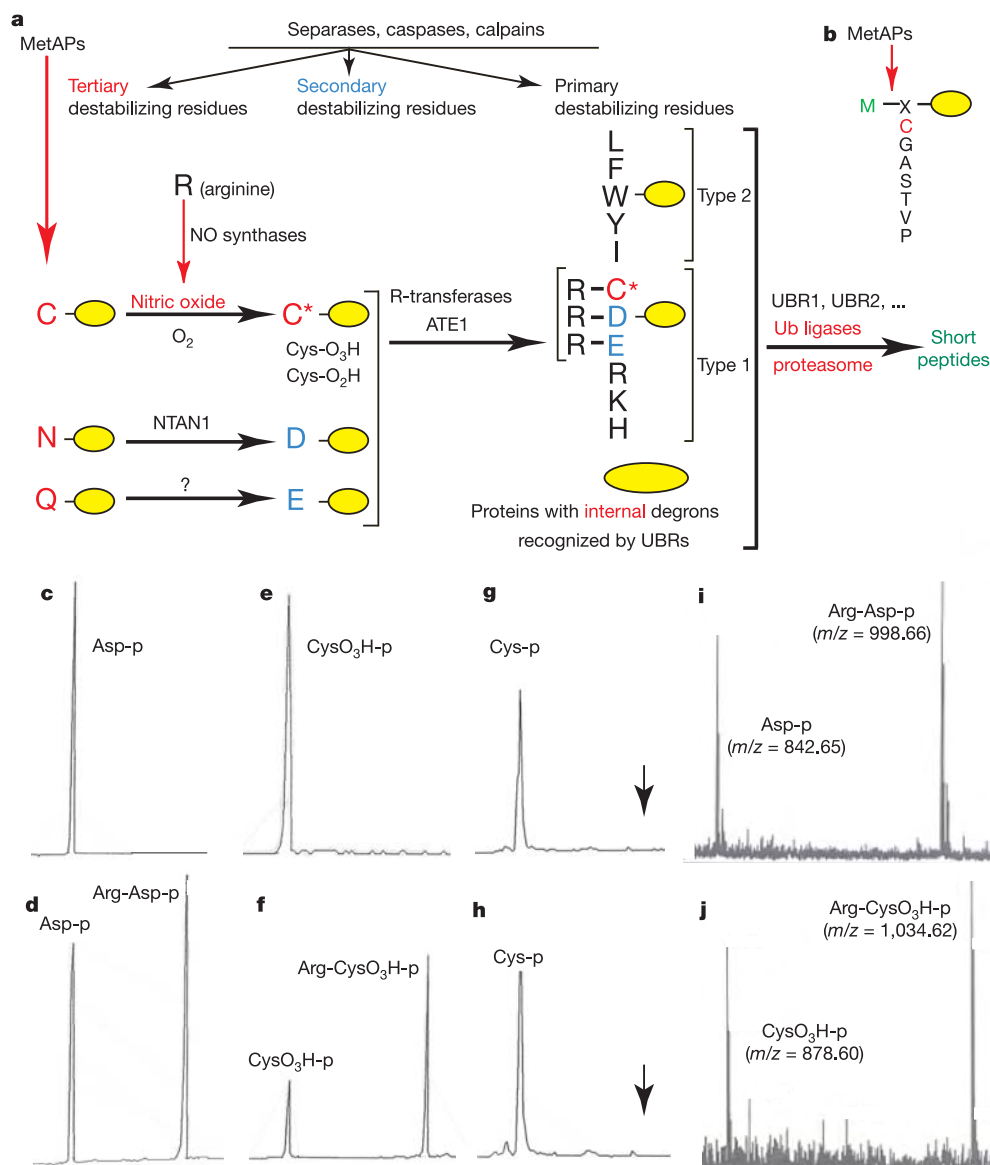


Figure 1 | N-terminal cysteine must be oxidized before its arginylation.

a, The mammalian N-end rule pathway. N-terminal residues are indicated by single-letter abbreviations for amino acids. Yellow ovals denote the rest of a protein substrate. The 'cysteine' (Cys) sector, in the upper left corner, describes the main discovery of this work: the NO-mediated oxidation of N-terminal Cys, with subsequent arginylation of oxidized Cys by *ATE1*-encoded isoforms of R-transferase. C* denotes oxidized Cys, either Cys-sulphinic acid (CysO₂H) or Cys-sulphonic acid (CysO₃H). Type 1 and type 2 primary destabilizing N-terminal residues are recognized by E3 ubiquitin (Ub) ligases of the N-end rule pathway, including UBR1 and UBR2. Through their other substrate-binding sites these E3 ubiquitin ligases also recognize internal (non-N-terminal) degrons in other substrates of the N-end rule pathway, denoted by a larger yellow oval. **b**, MetAPs remove Met from the N terminus of a polypeptide if the residue at position 2 belongs to the set of residues shown. **c–j**, N-terminal Cys must be oxidized before its

arginylation. Three eight-residue peptides are denoted as X-p; their N-terminal residues (X) were either Asp, Cys or CysO₃H. X-p was incubated with mouse *ATE1*-1 R-transferase at pH 7.5 in the presence of ATP, *S. cerevisiae* Arg-tRNA synthetase and tRNAs, followed by analyses of peptide products, either by capillary electrophoresis (CE) (**c–h**) or by MALDI-TOF MS (**i, j**). The x and y axes (not shown) in CE patterns correspond, respectively, to the time of elution from CE column and *A*₂₀₀. **c, d**, Arginylation assay with Asp-p, for 0 min (**c**) and 60 min (**d**), followed by CE. **e, f**, As in **c** and **d** but with CysO₃H-p. **g, h**, As in **c** and **d** but with Cys-p. Vertical arrows in **g** and **h** indicate the electrophoretic position of the (separately run) marker Arg-Cys-p, a chemically synthesized arginylated Cys-p. **i, j**, MALDI-TOF MS of the sample in **d, j**, MALDI-TOF MS of the sample in **f**. The molecular masses in **i** and **j** are of ionized [$+H^+$] derivatives of the molecules indicated on these panels in their un-ionized form.

distinguished from oxidation of Cys after its arginylation. Three eight-residue XHGSAGWL peptides (where X represents Asp, Cys or CysO₃H) were incubated with mouse ATE1-1 R-transferase¹¹ in the presence of ATP, *S. cerevisiae* Arg-tRNA synthetase and tRNAs. Both the Asp-peptide and the CysO₃H-peptide were efficiently arginylated, whereas the Cys-peptide was not arginylated, as determined by capillary electrophoresis (compare Fig. 1c–f with Fig. 1g, h), and confirmed, with regard to the identities of products, by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) MS (Fig. 1i, j). The same results were obtained with *S. cerevisiae* ATE1 R-transferase³³, a strong ‘sequelogue’³⁴ of mammalian R-transferases¹¹ (Supplementary Fig. S1). We conclude that the oxidation of N-terminal Cys is essential for its arginylation (Fig. 1c–j).

Increased levels of RGS proteins in *ATE1*^{−/−} embryos

RGS4 is a GTPase-activating protein (GAP) for specific Gα subunits of G proteins and is a member of the family of RGS proteins that downregulate signalling by G proteins^{30–32,35–37}. Our earlier work showed that RGS4 (which begins, as a nascent protein, with Met-Cys)

was a substrate of the N-end rule pathway in reticulocyte extract³⁸. The levels of endogenous RGS4 were also increased by proteasome inhibitors³⁷ but the pathway of RGS4 degradation *in vivo* remained to be identified.

We compared the levels of endogenous RGS4 between 12.5-day-old (E12.5) wild-type mouse embryos and *ATE1*^{−/−} embryos, which lacked R-transferases and therefore lacked arginylation¹². The level of RGS4 in E12.5 *ATE1*^{−/−} embryos was strikingly higher than in E12.5 wild-type embryos (Fig. 2a). Similar results were obtained in pairwise comparisons of RGS4 in specific tissues from E14.5 wild-type and *ATE1*^{−/−} embryos (Fig. 2b). Two other members of the RGS family, RGS5 and RGS16, also bear N-terminal Cys. Experiments analogous to those with RGS4 revealed strong increases in RGS5 and RGS16 levels in *ATE1*^{−/−} embryos (Fig. 2c–f). These protein patterns (Fig. 2a–f) were not caused by increased levels of the corresponding

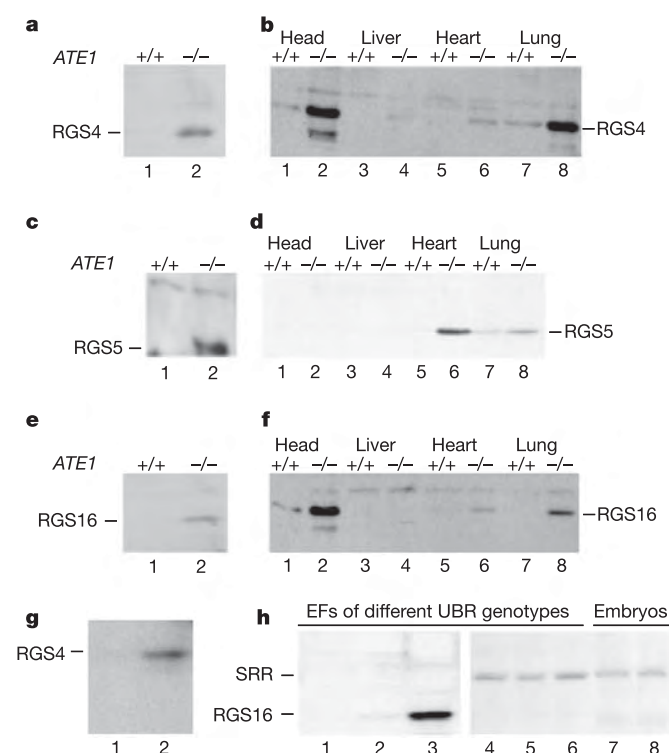


Figure 2 | Strongly increased levels of RGS4, RGS5 and RGS16 proteins in *ATE1*^{−/−} embryos. **a**, Lanes 1 and 2, equal amounts of total protein in extracts from wild-type (+/+) and *ATE1*^{−/−} E12.5 embryos were fractionated by 12% SDS–PAGE, followed by immunoblotting with anti-RGS4 antibody. **b**, As in **a** but with extracts from the indicated tissues of wild-type and *ATE1*^{−/−} E14.5 embryos. **c**, As in **a** but with anti-RGS5 antibody. **d**, As in **b** but with anti-RGS5 antibody. **e**, As in **a** but with anti-RGS16 antibody. **f**, As in **b** but with anti-RGS16 antibody. **g**, 3T3^{toff}RGS4_{th} cells expressing RGS4_{th} were grown in the presence of either ambient air (lane 1) or low (0.5%) oxygen concentration (lane 2). Equal amounts of total protein in extracts were subjected to immunoblotting with antibody against RGS4. **h**, Lanes 1–3, equal amounts of total protein in extracts from wild-type (lane 1), [*UBR1*^{−/−} *UBR2*^{−/−}] (lane 2) and *UBR1*^{ΔdnR2} (lane 3) cell lines were subjected to immunoblotting with antibody against RGS16. Lanes 4–6, as in lanes 1–3, respectively, but immunoblotting with antibody against serine racemase (SRR). Lanes 7, 8, as in lanes 4–6 but with extracts from wild-type (lane 7) and *ATE1*^{−/−} (lane 8) E12.5 embryos. The bands of 23K RGS4 (apparent molecular mass ~28 kDa), 20K RGS5 (apparent molecular mass ~21 kDa), 23K RGS16 (apparent molecular mass ~25 kDa) and 37K SRR are indicated in **a**–**h**.

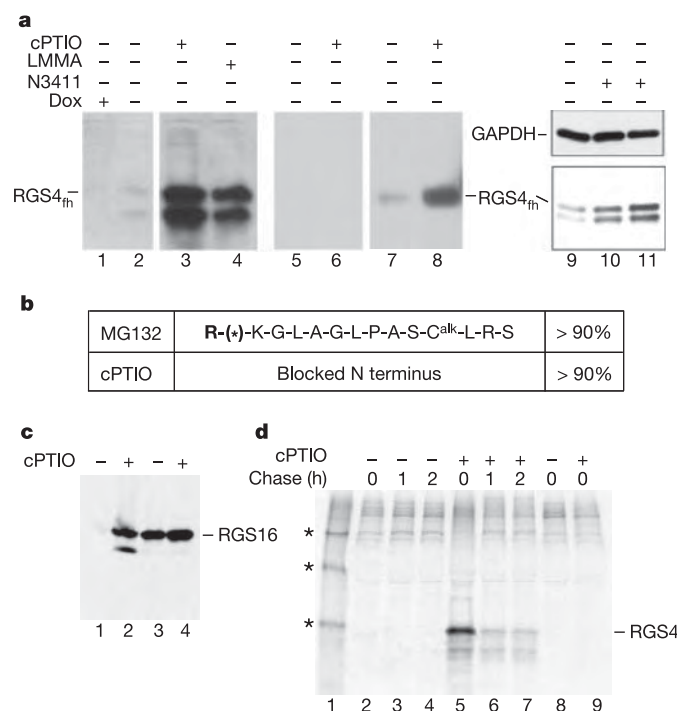


Figure 3 | Decreasing NO concentration *in vivo* stabilizes RGS4 and RGS16. **a**, 3T3^{toff}RGS4_{th} cells, expressing RGS4_{th}, were either untreated or treated with compounds that decrease the levels of intracellular NO, followed by immunoblotting with antibody against RGS4. Lanes 7–8 and 9–11 show two sets of experiments independent of those in lanes 1–4, with independently grown 3T3^{toff}RGS4_{th} cells. The concentrations of N3411 used in lanes 10 and 11 were 0.5 and 1 mM, respectively. Also indicated is the band of 37-kDa glyceraldehyde-3-phosphate dehydrogenase (GAPDH), a loading control, detected on the same membrane with antibody against GAPDH. **b**, Determination, through Edman degradation, of the N-terminal sequence of RGS4_{th} isolated from 3T3^{toff}RGS4_{th} cells that had been treated as described in the panel and in the text. C^{alk}, alkylated Cys residue. The asterisk after the N-terminal, post-translationally conjugated Arg (R) residue indicates a ‘sequenceable’ but unidentified residue, at the position of oxidized Cys residue in RGS4, in contrast to alkylated (identifiable) Cys (C^{alk}) at position 12. **c**, Lanes 1 and 2, immunoblotting, with antibody against RGS16, of extracts from NIH-3T3 cells that were either untreated (lane 1) or treated with cPTIO (lane 2). Lanes 3 and 4, as in lanes 1 and 2 but with *ATE1*^{−/−} EF cells¹² lacking arginylation. **d**, Lane 1, ¹⁴C-labelled protein markers, of molecular masses 66, 45 and 30 kDa (asterisks). Lanes 2–4, RGS4_{th}-expressing 3T3^{toff}RGS4_{th} cells were labelled for 10 min with ³⁵S-methionine/cysteine and chased for 1 and 2 h, followed by immunoprecipitation of extracts with antibody against RGS4, then by SDS–PAGE and autoradiography. Lanes 5–7, as in lanes 2–4 but with cPTIO-treated 3T3^{toff}RGS4_{th} cells. Lanes 8 and 9, as in lanes 2 and 5, respectively, but with 3T3^{toff} (RGS4_{th}-lacking) cells.

messenger RNAs, as indicated by both complementary DNA microarray comparisons (J.S., R.G.H., S. Choi., M. Simon and A.V., unpublished observations) and northern analyses (Supplementary Fig. S2).

To address the above issues differently, we employed mouse cells termed UBR1/2^{dnR2}. These [UBR1^{-/-}UBR2^{-/-}] cells, lacking two of several E3 ubiquitin ligases of the N-end rule pathway^{4,14}, stably express UBR2¹⁰⁴¹, the N-terminal fragment of mouse UBR2 that functions as a dominant-negative inhibitor of the (residual) N-end rule pathway in parental [UBR1^{-/-}UBR2^{-/-}] cells (J.S., Z. Xia, Y. T. Kwon and A.V., unpublished observations). The level of endogenous RGS16 was negligible in wild-type cells, barely detectable in [UBR1^{-/-}UBR2^{-/-}] cells and much higher in UBR1/2^{dnR2} cells (Fig. 2h, lanes 1–3). These results indicated yet again, in a setting independent from that of ATE1^{-/-} cells, that RGS16 was a substrate of the N-end rule pathway *in vivo*.

Decreasing NO *in vivo* stabilizes RGS4, RGS5 and RGS16

To examine the possibility that the oxidation of N-terminal Cys might involve NO, we constructed a mouse cell line, termed 3T3^{toff}RGS4_{fl}, that expressed RGS4-Flag-His₆ (RGS4_{fl}) from a doxycycline-repressible promoter. Treatment of 3T3^{toff}RGS4_{fl} cells with N^G-monomethyl-L-arginine (LMA), an inhibitor of NO synthases (NOSs), markedly increased the level of RGS4 *in vivo* (Fig. 3a, compare lanes 2 and 4). When 2-(4-carboxyphenyl)-4,4,5,5-tetramethylimidazole-1-oxyl-3-oxide (cPTIO), a cell-penetrating NO scavenger, was used to decrease NO levels, the increase in RGS4 was even more striking (Fig. 3a, compare lanes 2 and 3, and lanes 7 and 8). In the experiments of Fig. 3a, lanes 1–4 and 9–11, anti-RGS4 antibody detected two bands: the upper band's position was that expected for RGS4_{fl}; the lower was apparently a proteolytic fragment of RGS4_{fl}, because changes in its levels paralleled those of full-length RGS4_{fl}. The lower band was not observed in an otherwise identical but independent experiment of Fig. 3a, lanes 7 and 8. In addition, the same anti-RGS4 antibody did not detect RGS4 in parental 3T3^{toff} cells (Fig. 3a, lanes 5 and 6), and detected one RGS4 band in ATE1^{-/-} embryos (Fig. 2a).

Reverse-transcriptase-mediated polymerase chain reaction showed that 3T3^{toff}RGS4_{fl} cells expressed NOS1 (neuronal NOS) mRNA but little if any NOS2 (inducible NOS) and NOS3 (endothelial NOS) mRNAs (Supplementary Fig. S3). Given these findings, we also used an NOS inhibitor (N3411) that is highly selective for NOS1. In agreement with other NO results (Fig. 3a, lanes 1–8), and despite the near-confinement of inhibition by N3411 to NOS1, this

inhibitor substantially increased the levels of RGS4 (Fig. 3a, lanes 9–11). We also examined whether RGS4 was elevated in NOS1^{-/-} adult mice²⁴. Despite the presence of NOS2 and NOS3 in these mice, RGS4 was strongly increased in the NOS1^{-/-} lung, relative to the wild-type lung, and to a smaller extent in other tissues as well (Supplementary Fig. S4).

To determine whether a decrease in NO concentration could affect the arginylation of RGS4 *in vivo*, we used N-terminal (Edman) sequencing. RGS4_{fl} purified from control 3T3^{toff}RGS4_{fl} cells was completely or nearly completely (more than 90%) arginylated *in vivo* (Fig. 3b). In contrast, most RGS4_{fl} from cells that had been treated with cPTIO could not be sequenced by Edman degradation (Fig. 3b). The identity of the blocking modification of RGS4_{fl} remains to be determined. Thus, RGS4_{fl} from cells with decreased NO was largely unarginylated (Fig. 3b) and therefore not a target of the N-end rule pathway, in agreement with other results (Fig. 3a). In a pulse-chase assay, a negligible amount of ³⁵S-labelled RGS4_{fl} was present at the end of a 10-min pulse in untreated 3T3^{toff}RGS4_{fl} cells (Fig. 3d, lanes 2–4). In contrast, a strongly labelled band of RGS4_{fl} was observed with cPTIO-treated cells at the end of a 10-min pulse (Fig. 3d, compare lanes 2 and 5). Although stabilized in cPTIO-treated cells, RGS4_{fl} remained partly unstable even in the presence of cPTIO (Fig. 3d, lanes 5–7), which is consistent with the incomplete elimination of NO by cPTIO (refs 21, 22 and references therein).

RGS16 (but not RGS4) was naturally expressed in NIH-3T3 cells and in EF cell lines. Anti-RGS16 antibody detected trace amounts of RGS16 in untreated 3T3 cells but a much higher level of RGS16 after treatment with cPTIO (Fig. 3c, compare lanes 1 and 2). Crucially, this effect of cPTIO was nearly absent when the same experiment was performed with ATE1^{-/-} EF cells¹², which lacked arginylation (Fig. 3c, compare lanes 3 and 4), yielding independent evidence that the degradation of RGS16 requires both NO and arginylation. We also found that the level of RGS4 was much higher in 3T3^{toff}RGS4_{fl} cells grown under conditions of low (0.5%) oxygen concentration (Fig. 2g; see Discussion).

NO-dependent arginylation of RGS4 *in vitro*

Cys-RGS4, and also the otherwise identical Asp-RGS4 and Val-RGS4 (Fig. 1a), were produced in *Escherichia coli* with the ubiquitin fusion technique². X-RGS4 proteins were incubated with mouse ATE1-1 R-transferase in the presence of ATP, [³H]arginine, tRNA and other components of the analogous assay with eight-residue peptides (Fig. 1c–j), except that arginylation of X-RGS4 was detected by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and fluorography. In addition to untreated controls, X-RGS4 proteins were also preincubated with one of two NO donors, either diglutathionyl-dinitroso-iron (DNIC-(GSH)₂, a physiologically relevant NO carrier^{39,40}) or diethylenetriamine/nitric oxide adduct (DETA-NO). The results indicated an almost complete dependence of the arginylation of Cys-RGS4 *in vitro* on its previous exposure to a donor of NO (Fig. 4, compare lane 4 with lanes 5 and 6). (Dissolved oxygen and other gases were at levels that are normally present in buffers.) In contrast, Asp-RGS4 was efficiently arginylated irrespective of pretreatment with NO, and Val-RGS4 was not arginylated (Fig. 4). Thus, the NO dependence of *in vivo* arginylation of proteins bearing N-terminal Cys (Figs 2 and 3) can be reconstituted in an *in vitro* system (Fig. 4).

Sequence motif of NO-dependent N-degron

The advances described above (Figs 1–4) accounted for the following, previously unexplained, observation: the conversion of the position 3 residue of RGS4 (position 2 after Met removal) from basic Lys to uncharged Ser stabilized the resulting RGS4_{K3S} against degradation in a reticulocyte extract³⁸. In contrast, RGS4_{K3R}, in which Lys 2 was replaced by Arg, remained short-lived³⁸. Moreover, in contrast to wild-type RGS4, which was arginylated and bore CysO₃H at position 2 (Fig. 3d and ref. 12), RGS4_{K3S} that was transiently expressed in

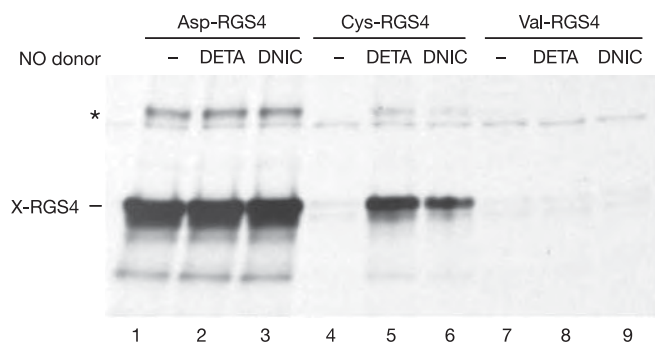


Figure 4 | *In vitro* reconstitution of nitric oxide-dependent arginylation of RGS4. Purified Asp-RGS4 (lanes 1–3), Cys-RGS4 (lanes 4–6) and Val-RGS4 (lanes 7–9) were incubated with [³H]arginine under conditions of the arginylation assay described in the text, followed by SDS-PAGE and fluorography. Lanes 1, 4 and 7, no pretreatment of X-RGS4 proteins. Lanes 2, 5 and 8, pretreatment of X-RGS4 proteins with DETA-NO. Lanes 3, 6 and 9, pretreatment of X-RGS4 proteins with DNIC-(GSH)₂. The asterisk denotes a minor ³H-labelled species whose apparent molecular mass and relative levels indicate that it might be a dimer of X-RGS4.

mouse cells was found to have a blocked N terminus (I. Davydov and A.V., unpublished observations). These results can now be explained, because S-nitrosylation of internal (non-N-terminal) Cys by NO in polypeptides depends, in particular, on the presence of a Cys-proximal basic residue, which facilitates the abstraction of H⁺ from the cysteine's thiol group^{21,22}. In agreement with this explanation, the second residues of MetAP-processed RGS4, RGS5 and RGS16 are Lys, Lys and Arg, respectively. In a different examination of the consensus motif (N-terminal Cys-(basic residue)), we used an antibody against serine racemase²⁰, a protein with a Cys-Ala N-terminal sequence, to examine whether the levels of racemase were increased in *ATE1*^{-/-} embryos or UBR1/2^{dnR2} cells. In contrast to results with RGS4, RGS5 and RGS16, no significant changes in racemase levels were observed (Fig. 2h, lanes 4–8), as would be expected from the requirement for a basic residue at position 2 for efficacious NO-mediated oxidation of N-terminal Cys.

Discussion

We have shown here that the oxidation of N-terminal Cys in a polypeptide is essential for the arginylation of Cys by *ATE1*-encoded R-transferases (Fig. 1c–j). We also discovered that the arginylation branch of the N-end rule pathway (Fig. 1a) is a sensor of nitric oxide (NO) that functions through its ability to destroy specific regulatory proteins bearing N-terminal Cys, at rates controlled by NO and apparently by oxygen as well (Figs 2–4). The first examples of these regulators, RGS4, RGS5 and RGS16, are also the first physiological substrates of mammalian N-end rule pathway. These proteins down-regulate specific α subunits of G proteins by increasing their GTPase activity^{30–32,35–37}. Through the conditional destruction of RGS4, RGS5 and RGS16, the N-end rule pathway is therefore involved in the regulation of signalling by G-protein-coupled receptors.

The above 'unification' of a ubiquitin-dependent proteolytic pathway and NO signalling opens up new vistas for understanding both. Most of the previously known regulation by NO was based on changes in the functional (for example enzymatic) activity of NO-modified proteins^{19–23,27–29}. In contrast, a Cys-containing N-degron of a protein makes possible the NO-mediated control of circuits that contain this protein, through its NO-dependent degradation by the N-end rule pathway. The observed stabilization of RGS4 in cells grown at low oxygen concentration (Fig. 2g) suggests a possible involvement of oxygen or its derivatives in this NO-dependent regulation. That would also be expected from the NO results alone (Figs 3 and 4), given multiple links between the chemistries of NO and oxygen *in vivo*^{19–23}. Our assay *in vitro* for NO-dependent arginylation of N-terminal Cys (Fig. 4) should eventually yield a detailed understanding of chemical transformations that result in oxidation of this uniquely positioned Cys residue.

The pathway of control by NO (Fig. 1a) targets proteins that bear N-terminal Cys followed by a basic residue. This motif is present in about 30 proteins encoded by the mouse (and human) genome, including RGS4, RGS5 and RGS16. More than half of non-RGS proteins in this set have entirely unknown functions, and the rest are barely characterized. The N-terminal Cys residues of RGS4, RGS5 and RGS16 can also be modified through palmitoylation³⁶. The two modifications are expected to be mutually exclusive and also act in opposite ways, in that palmitoylation increases the activity of RGS proteins as downregulators of G proteins³⁶, whereas NO-dependent arginylation and destruction of RGSs reduce their activity by decreasing their levels in a cell.

RGS4 is a physiological inhibitor of angiogenesis and other tubulogenesis pathways³². Upregulation of RGS4 perturbs cardiovascular homeostasis in mice³¹ and is a molecular correlate of human heart failure³⁰. Mouse *ATE1*^{-/-} embryos, which lack R-transferases and therefore lack arginylation, die before E17 with cardiovascular defects¹². Our findings that the levels of RGS4, RGS5 and RGS16 are greatly increased in the hearts and other organs of *ATE1*^{-/-} embryos

(Fig. 2), and that NO is required for the proteolytic downregulation of these RGSs (Figs 3 and 4), are likely to account in part for the known role of NO, at physiologically optimal levels, in suppressing pathological changes in the heart^{25,26,41}. The functions of NO in cardiovascular homeostasis include the stimulation of cyclic GMP formation by guanylyl cyclase and the regulation of cardiac contractility through S-nitrosylation of the calcium release channel^{21,23,41}. Our results revealed an entirely different, mutually nonexclusive mechanism of NO signalling in the heart and other organs: the control of regulatory proteins bearing N-terminal Cys through their NO-dependent, arginylation-mediated degradation by the N-end rule pathway. Thus, pharmacological manipulation of activities or expression of R-transferases may provide an alternative, more selective route to clinically beneficial effects that are currently achieved through drugs that alter the levels of NO.

Mammalian R-transferases are strong sequelogues³⁴ of yeast (fungal) *ATE1* R-transferases. However, whereas the inactivation of mouse *ATE1* results in embryonic lethality¹², a deletion of *S. cerevisiae ATE1* renders cells unable to degrade reporters with N-terminal Asp or Glu but has not been found to cause any other abnormal phenotype^{2,33}. Our findings indicate that one function of arginylation in this and other organisms might be to serve as a sensor of nitrosative/oxidative stress (Supplementary Information). It remains to be determined whether the discovered signalling by NO proceeds exclusively through the oxidation of Cys-containing N-degrons (Fig. 1a) or whether NO can also function at other steps of the N-end rule pathway, for example through S-nitrosylation of its ubiquitin ligases or R-transferases.

A preponderance of circuits relevant to the findings of this work involve arginine: first, Arg is a direct precursor of NO (Fig. 1a); second, the levels of Arg are tightly controlled in a cell and are often downregulated by invading pathogens; third, Arg is a part of Arg-tRNA (a co-substrate of R-transferase), indicating a possible connection between the N-end rule pathway and regulation of translation; fourth, Arg is a primary destabilizing residue and is also conjugated to N-end rule substrates bearing N-terminal Asp, Glu or oxidized Cys (Fig. 1a); and last, some Arg residues in proteins undergo methylation or deimination, the latter a conversion of positively charged Arg to uncharged citrulline^{42,43}. It remains to be determined whether methylation or deimination of Arg *in vivo* involves N-terminal Arg, and whether a set of circuits that has now been shown to connect the N-end rule pathway and the signalling by NO holds yet another Arg-linked surprise.

METHODS

Full technical details are provided in Supplementary Information.

Mouse embryos, immunoblotting and northern hybridization. *ATE1*^{-/-} and wild-type embryos were produced as described¹². E12.5 and E14.5 whole embryos or their organs were collected, followed by immunoblotting of embryo extracts with antibodies against RGS4 (a gift from S. Mumby), RGS16 (a gift from C. K. Chen and M. Simon), RGS5 (a gift from M. Greenwood), serine racemase (BD Biosciences), and mouse *ATE1*-1 R-transferase. The latter antibody was raised in rabbits against purified *ATE1*-1.

3T3^{toff}RGS4_{fl}, other cell lines, and treatments. Treatments of cells with cPTIO, with LMMA or with N3411 (L-N^ω-nitroarginine-2,4-L-diaminobutyric amide) were performed as described in Supplementary Information, followed by the preparation of extracts, SDS-PAGE and immunoblotting. The mouse 3T3^{toff}RGS4_{fl} cell line, which expressed RGS4-Flag-His₆ (RGS4_{fl}) from a doxycycline-repressible promoter, was constructed with MEF/3T3 'Tet-off' cells (BD Biosciences). Other mouse cell lines were UBR1/2^{dnR2} (see the text; J.S., Z. Xia, Y. T. Kwon and A.V., unpublished observations), *ATE1*^{-/-} EF cells and NIH-3T3 cells. Pulse-chase experiments were performed with 3T3^{toff}RGS4_{fl} cells, essentially as described¹².

Reporter peptides and recombinant proteins. The eight-residue Asp-peptide (DHGSGAWL) and Cys-peptide (CHGSGAWL) were synthesized by standard methods. CysO₃H-peptide was synthesized as described in Supplementary Information. *S. cerevisiae* Arg-tRNA synthetase (RRS1) was expressed in *E. coli* and purified by chromatography. *S. cerevisiae ATE1* R-transferase³³ was carboxy-terminally tagged with His₆, expressed in *E. coli* and purified

with Ni^{2+} -nitrilotriacetate (NTA)-agarose. Mouse ATE1-1 R-transferase^{11,12} was C-terminally epitope-tagged, expressed with a baculovirus and purified with Ni^{2+} -NTA-agarose. X-RGS4 proteins were produced and purified as described in Supplementary Information.

Arginylation assays, NO donors, capillary electrophoresis, and other assays.

A sample for performing the arginylation of a reporter peptide contained, in addition to reaction buffer, one of three eight-residue synthetic peptides, a mixture of *S. cerevisiae* tRNAs, purified *S. cerevisiae* Arg-tRNA synthetase, and either purified mouse ATE1-1 R-transferase or purified *S. cerevisiae* ATE1 R-transferase. Reactions were stopped by the addition of trifluoroacetic acid, followed by analyses of peptides by either capillary electrophoresis or MALDI-TOF MS. Details of the arginylation assay with [³H]arginine, purified X-RGS4 proteins and the NO donors DNIC-[GSH]₂ and DETA-NO are described in Supplementary Information.

For N-terminal sequencing by Edman degradation, RGS4_{th} was isolated from 3T3^{off}RGS4_{th} cells, alkylated with iodoacetamide, fractionated by 12% SDS-PAGE and processed for sequencing¹¹.

Received 26 May; accepted 15 July 2005.

- Bachmair, A., Finley, D. & Varshavsky, A. *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 (1986).
- Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl Acad. Sci. USA* **93**, 12142–12149 (1996).
- Varshavsky, A. The N-end rule and regulation of apoptosis. *Nature Cell Biol.* **5**, 373–376 (2003).
- Kwon, Y. T. *et al.* Female lethality and apoptosis of spermatocytes in mice lacking the UBR2 ubiquitin ligase of the N-end rule pathway. *Mol. Cell. Biol.* **23**, 8255–8271 (2003).
- Bachmair, A. & Varshavsky, A. The degradation signal in a short-lived protein. *Cell* **56**, 1019–1032 (1989).
- Suzuki, T. & Varshavsky, A. Degradation signals in the lysine-asparagine sequence space. *EMBO J.* **18**, 6017–6026 (1999).
- Hershko, A., Ciechanover, A. & Varshavsky, A. The ubiquitin system. *Nature Med.* **10**, 1073–1081 (2000).
- Pickart, C. Back to the future with ubiquitin. *Cell* **116**, 181–190 (2004).
- Baker, R. T. & Varshavsky, A. Yeast N-terminal amidase: a new enzyme and component of the N-end rule pathway. *J. Biol. Chem.* **270**, 12065–12074 (1995).
- Kwon, Y. T. *et al.* Altered activity, social behaviour, and spatial memory in mice lacking the NTAN1 amidase and the asparagine branch of the N-end rule pathway. *Mol. Cell. Biol.* **20**, 4135–4148 (2000).
- Kwon, Y. T., Kashina, A. S. & Varshavsky, A. Alternative splicing results in differential expression, activity, and localization of the two forms of arginyl-tRNA-protein transferase, a component of the N-end rule pathway. *Mol. Cell. Biol.* **19**, 182–193 (1999).
- Kwon, Y. T. *et al.* An essential role of N-terminal arginylation in cardiovascular development. *Science* **297**, 96–99 (2002).
- Du, F., Navarro-Garcia, F., Xia, Z., Tasaki, T. & Varshavsky, A. Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. *Proc. Natl Acad. Sci. USA* **99**, 14110–14115 (2002).
- Tasaki, T. *et al.* A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol. Cell. Biol.* **25**, 7120–7136 (2005).
- Gonda, D. K. *et al.* Universality and structure of the N-end rule. *J. Biol. Chem.* **264**, 16700–16712 (1989).
- Turner, G. C., Du, F. & Varshavsky, A. Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* **405**, 579–583 (2000).
- Ditzel, M. *et al.* Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nature Cell Biol.* **5**, 467–473 (2003).
- Rao, H., Uhlmann, F., Nasmyth, K. & Varshavsky, A. Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* **410**, 955–960 (2001).
- Ignarro, L. J. Nitric oxide as a unique signalling molecule in the vascular system: a historical overview. *J. Physiol. Pharmacol.* **53**, 503–514 (2002).
- Boehning, D. & Snyder, S. H. Novel neural modulators. *Annu. Rev. Neurosci.* **26**, 105–131 (2003).
- Hess, D. T., Matsumoto, A., Kim, S.-O., Marshall, H. E. & Stamler, J. S. Protein S-nitrosylation: purview and parameters. *Nature Rev. Mol. Cell Biol.* **6**, 150–166 (2005).
- Nathan, C. Specificity of a third kind: reactive oxygen and nitrogen intermediates in cell signalling. *J. Clin. Invest.* **111**, 769–778 (2003).
- Eu, J. P., Sun, J., Xu, L., Stamler, J. S. & Meissner, G. The skeletal muscle calcium release channel: coupled O₂ sensor and NO signalling functions. *Cell* **102**, 499–509 (2000).
- Packer, M. A. *et al.* Nitric oxide negatively regulates mammalian adult neurogenesis. *Proc. Natl Acad. Sci. USA* **100**, 9566–9571 (2003).
- Feng, Q. *et al.* Development of heart failure and congenital septal defects in mice lacking endothelial nitric oxide synthase. *Circulation* **106**, 873–879 (2002).
- Barouch, L. A. *et al.* Nitric oxide regulates the heart by spatial confinement of nitric oxide synthase isoforms. *Nature* **416**, 337–340 (2002).
- van Coelln, R., Dawson, V. L. & Dawson, T. M. Parkin-associated Parkinson's disease. *Cell Tissue Res.* **318**, 175–184 (2004).
- Yao, D. *et al.* Nitrosative stress linked to sporadic Parkinson's disease: S-nitrosylation of parkin regulates its E3 ubiquitin ligase activity. *Proc. Natl Acad. Sci. USA* **101**, 10810–10814 (2004).
- Tanaka, K., Suzuki, T., Hattori, N. & Mizuno, Y. Ubiquitin, proteasome and parkin. *Biochim. Biophys. Acta* **1695**, 226–238 (2004).
- Wieland, T. & Mittman, C. Regulators of G-protein signalling: multifunctional proteins with impact on signalling in the cardiovascular system. *Pharmacol. Therapeut.* **97**, 95–115 (2003).
- Rogers, J. S. *et al.* RGS4 reduces contractile dysfunction and hypertrophic gene induction in Gαq-overexpressing mice. *J. Mol. Cell. Cardiol.* **33**, 209–218 (2001).
- Albig, A. R. & Schiemann, W. P. Identification and characterization of regulator of G protein signalling 4 (RGS4) as a novel inhibitor of tubulogenesis: RGS4 inhibits mitogen-activated protein kinases and vascular endothelial growth factor signalling. *Mol. Biol. Cell* **16**, 609–625 (2005).
- Balzi, E., Choder, M., Chen, W., Varshavsky, A. & Goffeau, A. Cloning and functional analysis of the arginyl-tRNA-protein transferase gene *ATE1* of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **265**, 7464–7471 (1990).
- Varshavsky, A. 'Spalog' and 'sequelog': neutral terms for spatial and sequence similarity. *Curr. Biol.* **14**, R181–R183 (2004).
- Berman, D. M. & Gilman, A. G. Mammalian RGS proteins: Barbarians at the gate. *J. Biol. Chem.* **273**, 1269–1272 (1998).
- Smotryns, J. E. & Linder, M. E. Palmitoylation of intracellular signalling proteins: regulation and function. *Annu. Rev. Biochem.* **73**, 559–587 (2004).
- Krumins, A. M. *et al.* Differentially regulated expression of endogenous RGS4 and RGS7. *J. Biol. Chem.* **279**, 2593–2599 (2004).
- Davydov, I. V. & Varshavsky, A. RGS4 is arginylated and degraded by the N-end rule pathway *in vitro*. *J. Biol. Chem.* **275**, 22931–22941 (2000).
- Mülsch, A., Lurie, D. J., Seimenis, I., Fichtlscherer, B. & Foster, M. A. Detection of nitrosyl-iron complexes by proton-electron-double-resonance imaging. DNIC as endogenous NO carrier. *Free Radic. Biol. Med.* **27**, 636–646 (1999).
- Becker, K., Savvides, S. N., Keese, M., Schirmer, R. H. & Karplus, P. A. Enzyme inactivation through sulfhydryl oxidation by physiologic NO-carriers. *Nature Struct. Biol.* **5**, 267–271 (1998).
- Kempf, T. & Wollert, K. C. Nitric oxide and the enigma of heart hypertrophy. *Bioessays* **26**, 608–615 (2004).
- Bedford, M. T. & Richard, S. Arginine methylation: an emerging regulator of protein function. *Mol. Cell* **18**, 263–272 (2005).
- Vossenaar, E. R., Zendman, A. J. W., van Venrooij, W. J. & Puij, G. J. M. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *BioEssays* **25**, 1106–1118 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Enikolopov, T. Michurina and J. M. Encinas for *NOS1^{-/-}* mice; A. Mülsch for DNIC-[GSH]₂; M. Shahgholi and J. Zhou for MS analyses of peptides; F. Rusnak and G. Hathaway for protein sequencing; J. Racs and S. Horvath for peptide synthesis; L. del Carmen Sandoval, B. W. Kennedy and S. Pease for advice and assistance with mouse mutants; Y. T. Kwon for [UBR1^{-/-}UBR2^{-/-}] cells; R. Baker for plasmids that enabled the USP2-based ubiquitin fusion technique; G. Eriani and F. Du for gifts of other plasmids; Z. Xia for USP2 enzyme; R. Roberts for use of his laboratory equipment; and E. Graciet and C. Brower for comments on the manuscript. Purification of mouse ATE1-1 was performed at CalTech's Protein Expression Center by I. K. Nangiana and the late P. Snow. We dedicate this paper to the memory of Dr Snow. This work was supported by grants from the NIH and the Ellison Medical Foundation to A.V.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.V. (avarsh@caltech.edu).

A large dust/ice ratio in the nucleus of comet 9P/Tempel 1

Michael Küppers¹, Ivano Bertini², Sonia Fornasier², Pedro J. Gutierrez³, Stubbe F. Hviid¹, Laurent Jorda⁴, Horst Uwe Keller¹, Jörg Knollenberg⁵, Detlef Koschny⁶, Rainer Kramm¹, Luisa-Maria Lara³, Holger Sierks¹, Nicolas Thomas⁷, Cesare Barbieri², Philippe Lamy⁴, Hans Rickman⁸, Rafael Rodrigo³ & the OSIRIS team*

Comets spend most of their life in a low-temperature environment far from the Sun. They are therefore relatively unprocessed and maintain information about the formation conditions of the planetary system, but the structure and composition of their nuclei are poorly understood. Although *in situ*¹ and remote² measurements have derived the global properties of some cometary nuclei, little is known about their interiors. The Deep Impact mission³ shot a projectile into comet 9P/Tempel 1 in order to investigate its interior. Here we report the water vapour content (1.5×10^{32} water molecules or 4.5×10^6 kg) and the cross-section of the dust (330 km^2 assuming an albedo of 0.1) created by the impact. The corresponding dust/ice mass ratio is probably larger than one, suggesting that comets are 'icy dirtballs' rather than 'dirty snowballs' as commonly believed⁴. High dust velocities (between 110 m s^{-1} and 300 m s^{-1}) imply acceleration in the comet's coma, probably by water molecules sublimated by solar radiation. We did not find evidence of enhanced activity of 9P/Tempel 1 in the days after the impact, suggesting that in general impacts of meteoroids are not the cause of cometary outbursts.

On 4 July 2005, the impactor of the NASA Deep Impact mission hit the surface of comet 9P/Tempel 1 with a relative velocity of 10.2 km s^{-1} . The collision of the impactor with a mass of 362 kg was expected to generate a crater (predicted diameter $\sim 100\text{--}125 \text{ m}$; ref. 3) and eject cometary material. Possibly it would also trigger an outburst and a new active area on the comet's surface.

The event was observed by many ground-based and some space-based telescopes. We used the scientific OSIRIS camera system on board the ESA Rosetta spacecraft to observe the comet before and after the impact. OSIRIS is composed of a narrow angle camera (NAC) and a wide angle camera (WAC). Both cameras are unobstructed mirror systems and are equipped with $2,048 \times 2,048$ pixel backlit/back-thinned CCD detectors with a sensitivity range from 245 nm to 1,000 nm and a maximum quantum efficiency of 95%. At the time of impact Rosetta was located at a distance of 0.53 AU from 9P/Tempel 1 with a solar elongation angle of 91° and a phase angle of 69° . The heliocentric distance of the comet was 1.51 AU. OSIRIS started to observe the comet on 28 June 2005 at 23:45 UTC and finished on 14 July 2005 15:00 UTC. The time resolution was generally 1.5–3 h and was increased (with the NAC) to approximately 1 min between 30 min before and 90 min after the impact. The NAC imaged the extended dust coma in different filters. The WAC was commanded to take images through the OH and a neighbouring continuum filter. In addition Na, CN, and O I filters were used to monitor further gas emissions.

An asymmetry of the ejected dust cloud is clearly visible for several days after the impact (Fig. 1). The analysis presented separates this debris from the background of the normal coma. The asymmetry of the gas (OH) is less visible because of the reduced spatial resolution ($31,200 \text{ km}$) and the lower signal to noise ratio. Immediately after impact both cameras act like photometers until the impact-related dust and gas leaves the corresponding resolution element.

The water (H_2O) production rate of comet 9P/Tempel 1 was derived from the OH emission (308 nm). A scaled image of the ultraviolet dust continuum (at 375 nm) was subtracted from each OH image, assuming solar type reflectivity for the cometary dust. Pre-launch laboratory calibration and observations of Vega (α Lyrae) were used for conversion of data numbers into flux units.

To estimate the water production rate before the impact, the flux from OH molecules was added within circular areas with radii between 31,200 km and 156,000 km centred on the position of the cometary nucleus. A fluorescence efficiency of $1.5 \times 10^{-22} \text{ W molecule}^{-1}$ at 1 AU (ref. 5) was scaled to the heliocentric distance of the comet and used to convert fluxes to number of molecules. The Haser formula^{6,7} for the number of molecules within a circular aperture is used to convert the number of OH molecules to the water production rate of the comet. The underlying assumptions are that OH is derived from dissociation of cometary water, and that both parent (H_2O) and daughter (OH) molecules are flowing radially away from the nucleus with a constant velocity. We used typical values of the photodissociation timescales for H_2O ($7.67 \times 10^4 \text{ s}$ at 1 AU) and OH ($1.32 \times 10^5 \text{ s}$)⁸, scaled to the heliocentric distance of comet 9P/Tempel 1. A fraction (86%) of all water molecules dissociate into $\text{OH} + \text{H}$ (ref. 9). The resulting pre-impact water production rate is $(3.4 \pm 0.5) \times 10^{27}$ molecules s^{-1} for a typical gas outflow velocity of 0.7 km s^{-1} at 1.5 AU from the Sun¹⁰, and $(5.8 \pm 1) \times 10^{27}$ molecules s^{-1} for an outflow velocity of 1 km s^{-1} , which is frequently used in the derivation of Haser parameters¹¹. The higher value for the production rate is in good agreement with other estimates from near-ultraviolet measurements¹².

Figure 2 shows the number of OH molecules created by photodissociation of H_2O in the impact cloud. We estimate that $(1.5 \pm 0.5) \times 10^{32}$ water molecules (or $4.6 \times 10^6 \text{ kg}$) were created by the impact. This value does not depend on the outflow velocity of the water or OH, and corresponds to approximately 20% of the water molecules that were present in the cometary coma owing to normal activity.

¹Max-Planck Institut für Sonnensystemforschung, Max-Planck-Str. 2, 37191 Katlenburg-Lindau, Germany. ²Dipartimento di Astronomia and CISAS, Università di Padova, Vicolo dell'Osservatorio 5, 35100 Padova, Italy. ³Instituto de Astrofísica de Andalucía-CSIC, C/ Camino Bajo de Huétor, 50, 18008 Granada, Spain. ⁴Laboratoire d'Astrophysique de Marseille, Traverse du Siphon, Les Trois Lucs BP 8, 13376 Marseille, France. ⁵DLR Institute for Planetary Research, Rutherfordstr. 2, 12489 Berlin, Germany. ⁶European Space Agency, ESTEC, SCI-SB, Keplerlaan 1, Postbus 299, 2200 AG Noordwijk, The Netherlands. ⁷Physikalisches Institut, Abteilung Weltraumforschung und Planetologie, Universität Bern, Sidlerstr. 5, 3012 Bern, Switzerland. ⁸Uppsala Astronomical Observatory, Box 515, 75120 Uppsala, Sweden.

*Lists of participants and affiliations appear at the end of the paper.

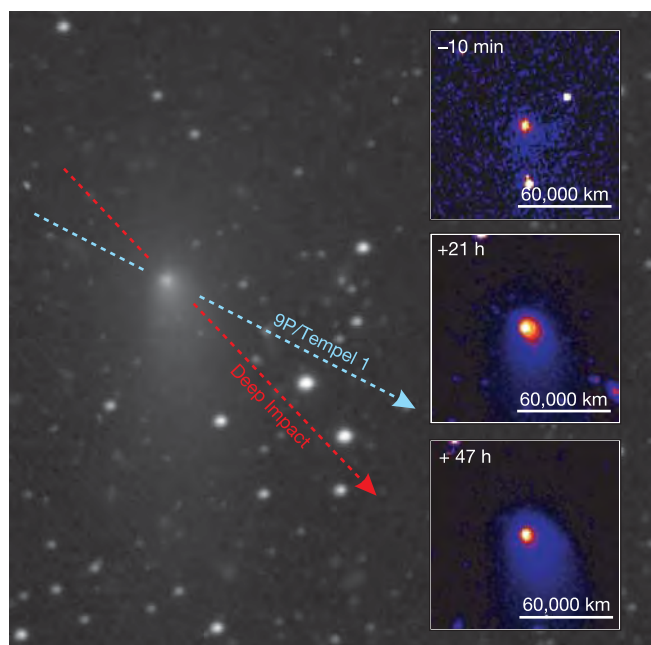


Figure 1 | The Comet 9P/Tempel 1 as seen by the Osiris NAC. The NAC focal length is 717 mm with a 87.5-mm aperture and a 2.35° field of view. The imaging scale at the comet was $1,500 \text{ km pixel}^{-1}$ (3,000 km image resolution). The images were taken through the orange filter (at $648 \pm 43 \text{ nm}$). The signal from the comet in the wavelength range of the orange filter is predominantly solar radiation reflected from cometary dust. The two dotted lines indicate the projected orbits of both the comet (blue) and the Deep Impact spacecraft (red) as seen from the Rosetta spacecraft. The insets show the dust coma and the expanding debris cloud at the time of impact, +21 h, and +47 h. In all cases the Sun is to the top. The outflow speed of the dust is estimated from the expansion speed of the cloud by analysing the brightness decrease within circular apertures between 1,500 km and 12,000 km radius. The derived speed is a lower limit to the actual outflow velocity of the dust because we measured only the velocity component parallel to the image plane. We estimated the speed of the dust as the ratio between the radius of an aperture and the time passed until the brightness in the aperture started to decrease. The so-determined speed of $300 \pm 50 \text{ m s}^{-1}$ refers to the fastest dust particles. If one uses the time at which the brightness of the impact-created dust cloud within an aperture has decreased to half of its maximum, the outflow speed of typical dust grains can be estimated. The corresponding number is $110 \pm 10 \text{ m s}^{-1}$. There is no significant change of these velocities between 1,500 km and 12,000 km from the nucleus.

During the first 40 min after the impact, the brightness of the cometary dust increased by a factor of 4.5 within an circular area (aperture) of 3,000 km (2 pixel) radius at the comet (Fig. 3 inset). During the following 50 min the brightness levelled off. The brightness increase lasted at least an order of magnitude longer than the expected crater formation time of 3–6 min (ref. 13). After 1.5 h the brightness in the 2-pixel aperture began slowly to decrease because the impact-created dust cloud started to move out of the selected aperture. In the following week, the brightness decreased slowly to its pre-impact level. The dust expands with a velocity (projected on the image plane) of typically 110 m s^{-1} (Fig. 1). The fastest dust, moving with a speed of at least 300 m s^{-1} , consists of tiny particles. The effects of solar radiation pressure are visible 2 d after impact (see Fig. 1).

The observed maximum brightness increase of $7.5 \times 10^{-7} \text{ W m}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$ in the orange filter, the difference between the peak and the pre-impact brightness (Fig. 3), corresponds to newly created dust particles with a reflecting surface of $330 \pm 30 \text{ km}^2$ assuming optically thin conditions and a scattering albedo of 0.1 (ref. 14). Very

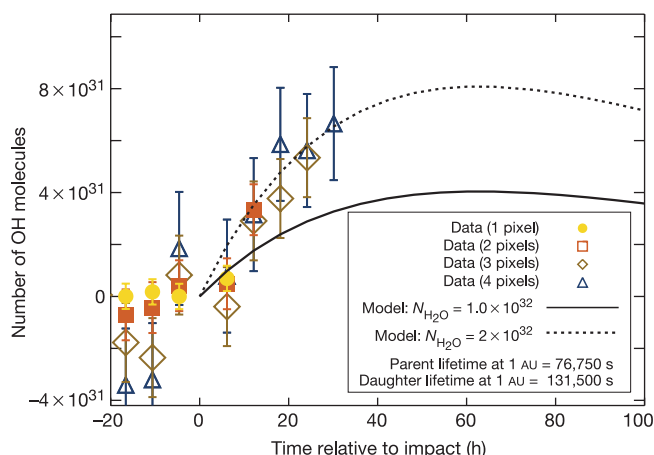


Figure 2 | Number of OH molecules created by the impact, as measured with the OSIRIS/WAC. The WAC focal length is 136 mm with a 25-mm aperture and a 12.1° field of view. The pixel scale is $7,800 \text{ km pixel}^{-1}$, but in the images used here the resolution was reduced to $31,200 \text{ km}$ by on-chip binning over 4×4 pixels. The number of OH molecules created by dissociation of water molecules liberated by the impact is derived from the difference between the total number of OH molecules measured at a certain time and within a circular area centred on the comet and the corresponding pre-impact number created by the normal activity of the comet. To convert the number of OH molecules to the number of impact-created water molecules, we monitor the number of OH molecules (the OH emission) in each circular area until the expanding cloud of molecules leaves that particular area. Then the number of H_2O molecules can be determined from

$$N_{\text{H}_2\text{O}}(t) = N_{\text{H}_2\text{O}}(t=0) \exp(-t/\tau_{\text{H}_2\text{O}})$$

and

$$N_{\text{OH}}(t) = f_{\text{OH}} \times \int_0^t -\frac{dN_{\text{H}_2\text{O}}}{dt'} \exp(-(t-t')/\tau_{\text{OH}}) dt' \\ = f_{\text{OH}} N_{\text{H}_2\text{O}}(t=0) \frac{\tau_{\text{OH}}}{\tau_{\text{OH}} - \tau_{\text{H}_2\text{O}}} (\exp(-t/\tau_{\text{OH}}) - \exp(-t/\tau_{\text{H}_2\text{O}}))$$

Here N is the number of molecules and τ the dissociation timescale. f_{OH} is the fraction of H_2O molecules that dissociated into $\text{OH} + \text{H}$. Data for apertures between 31,000 km and 125,000 km are shown, along with models that represent the minimum and maximum number of water molecules consistent with the data. The median number of molecules seen in the same aperture before the impact was subtracted from the data. The error bars contain the standard deviation of the number of molecules seen pre-impact and the statistical error of each measurement. They do not contain the $\sim 20\%$ uncertainty of the intensity calibration.

small particles (radius $\leq 1 \mu\text{m}$) scatter light most effectively per unit mass. If only $1\text{-}\mu\text{m}$ particles were present, their total mass was $4.4 \times 10^5 \text{ kg}$ for a dust density of $1,000 \text{ kg m}^{-3}$ (the density of individual dust grains is expected to be larger than that of the porous cometary nucleus). This is an unrealistic lower limit. If the size and mass distribution of the dust created by the impact is similar to that measured in the coma of comets (power law distribution with an exponent γ between -4 and -3.3 ; refs 15, 16), the dust/water ice ratio of the material excavated by the impact is one or larger. This ratio is sensitive to the exact mass exponent and to the cut-off mass at the large end. It is 1 for $\gamma = -3.8$, and a cut-off mass of 10,000 kg (ref. 17). An exponent of -3.6 (ref. 15) yields a dust/ice ratio of 20, and -3.5 a ratio of 100. For a dust/ice ratio of one, the total ejecta mass, water and dust, amounts to about 10^7 kg ; this corresponds to the volume of a bowl-shaped crater with a depth/diameter ratio of 1/4 and a radius of about 30 m (ref. 18). This is a lower limit to the crater size, because the dust/gas ratio is probably > 1 and part of the excavated material may not have left the comet. If γ is the typical range (> -3.8) and the excavated volume is larger, then the dust/ice

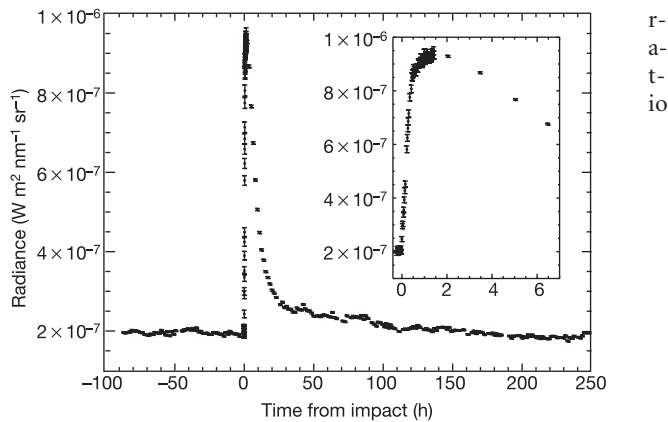


Figure 3 | Light curve of the cometary dust in the orange filter. The data show the brightness in a circular field of 3,000 km radius at the comet. Error bars represent the Poisson noise only, and do not include the $\sim 10\%$ uncertainty of the intensity calibration. For comparison, we estimate that the cometary nucleus with a cross-section of 33 km^2 and a geometric albedo of 0.05 (ref. 20) contributes about $7.5 \times 10^{-8} \text{ W m}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$, or 10% of the brightness of the impact-created dust cloud at the phase angle of Rosetta.

would be considerably larger than one, supporting the model of an icy dirtball¹⁹ rather than a dirty snowball.

Several processes could have caused the brightness increase after the impact: an originally opaque dust cloud may become transparent as it expands, the activity of the cometary nucleus may be enhanced to an outburst by the impact, or the dust particles (consisting of icy grains) in the cloud created by the impact may fragment owing to the sublimation of the water ice. The dust cloud expands to a distance of $\sim 264 \text{ km}$ from the nucleus (speed 110 m s^{-1}) within the 40 min of the initial steep brightness increase. The corresponding cross-section of $2 \times 10^5 \text{ km}^2$ is two orders of magnitude larger than the observed total scattering surface of the dust. Therefore optical depth plays a role in the brightness increase only during the first few minutes.

The new active area created by the impact is two orders of magnitude smaller than the fraction (10%, ref. 20) of the nucleus surface required for normal pre-impact activity. Even if the activity were driven by sublimation of the most volatile CO, a crater diameter of 400 m would be needed to double the production rate of the comet.

The most likely explanation for the brightness increase during the first 40 min after the impact is that the dust grains fragmented into smaller and smaller grains, increasing their scattering efficiency. The fragmentation is most probably caused by water ice sublimation driven by solar irradiation. A sphere of 0.2 mm radius of water ice with an albedo of 0.05 (caused by dust inclusions) sublimates in 40 min (ref. 21). Destruction will be faster in the more realistic case of an ice–dust mixture and a combination of sublimation and fragmentation.

The fragmentation of originally more coarse-grained dust created by the impact is consistent with size distributions of impact ejecta measured in the laboratory²². They are less steep than that derived for comets. The final size distribution of the dust grains will probably be derived from the shape of the dust cloud observed by Earth-based telescopes.

The kinetic energy of the impactor ($\sim 2 \times 10^{10} \text{ J}$) is insufficient to provide the energy required to sublimate the observed amount of water ($\sim 10^{13} \text{ J}$). Therefore the overwhelming part of water was ejected as icy grains. The energy for the sublimation of the grains and the acceleration of the dust in the coma is provided by sunlight.

The continuous observations of OSIRIS covering 17 d around the time of impact suggest the following scenario. The impactor created a crater of $\geq 30 \text{ m}$ radius. The crater ejecta consisted essentially of icy (water ice) dust grains. The present preliminary estimates are in agreement with a dust/water-ice mass ratio probably considerably larger than one. If a large crater was indeed formed, the dust/water-ice ratio would increase with the volume of the crater. The originally ejected grains were fragmented into fine micrometre-sized (or even smaller) grains by sublimation of the water ice due to solar irradiation. The initially optically thick ejecta cloud became transparent very fast, within the first few minutes after the impact. The fragmentation of the icy grains rapidly increased the brightness of the cloud. The final size distribution was reached after about 40 min, when the brightness increase levelled off. After that, the cloud continued to expand with speeds of more than 110 m s^{-1} (a fine dust component reached speeds of 300 m s^{-1}), accelerated beyond the ejecta speed by the sublimated water molecules. The brightness of the comet assumed its nominal level about a week after the impact. Any additional activity triggered by the impact was small compared to the normal activity level of comet 9P/Tempel 1, and outbursts triggered by internal (chemical) energy, for example, can be excluded. Therefore we conclude that impacts of metre-sized meteoroids are not the cause of frequently observed outbursts of comets. The OSIRIS observations suggest that the dust/ice ratio in comet 9P/Tempel 1 may be substantially larger than one.

Received 25 July; accepted 7 September 2005.

- Keller, H. U., Britt, D., Buratti, B. J. & Thomas, N. in *Comets II* (eds Festou, M. C., Keller, H. U. & Weaver, H. A.) 211–222 (Univ. Arizona Press, Tucson, 2004).
- Lamy, P. L., Toth, I., Fernández, Y. R. & Weaver, H. A. in *Comets II* (eds Festou, M. C., Keller, H. U. & Weaver, H. A.) 223–264 (Univ. Arizona Press, Tucson, 2004).
- A'Hearn, M. A., Belton, M. J. S., Delamere, A. & Blume, W. H. Deep Impact: A large-scale active experiment on a cometary nucleus. *Space Sci. Rev.* **117**, 1–21 (2005).
- Whipple, F. A comet model. I. The acceleration of Comet Encke. *Astrophys. J.* **111**, 375–394 (1950).
- Schleicher, D. G. & A'Hearn, M. F. The fluorescence of cometary OH. *Astrophys. J.* **331**, 1058–1077 (1988).
- Haser, L. Distribution d'intensité dans la tête d'une comète. *Bull. Acad. R. Sci. Liege* **43**, 740–750 (1957).
- Krishna Swamy, K. S. *Physics of Comets* 2nd ed. (World Scientific Series in Astronomy and Astrophysics, Vol. 2, World Scientific Publishing, Singapore, 1997).
- Budzien, S. A., Festou, M. C. & Feldman, P. D. Solar flux variability and the lifetimes of cometary H₂O and OH. *Icarus* **164**, 164–188 (1994).
- Huebner, W. F., Keady, J. J. & Lyon, S. P. Solar photo rates for planetary atmospheres and atmospheric pollutants—Photo rate coefficients and excess. *Astrophys. Space Sci.* **195**, 7–289 (1992).
- Combi, M. R., Harris, W. M. & Smyth, W. H. in *Comets II* (eds Festou, M. C., Keller, H. U. & Weaver, H. A.) 523–552 (Univ. Arizona Press, Tucson, 2004).
- A'Hearn, M. F. A., Millis, R. L., Schleicher, D. G., Osip, D. J. & Birch, P. V. The ensemble properties of comets: Results from narrowband photometry of 85 comets, 1976–1992. *Icarus* **118**, 223–270 (1995).
- Schleicher, D., Barnes, K. *IAU Circ.* No. 8546 (2005).
- Schultz, P. R., Ernst, C. M. & Anderson, J. L. B. Expectations for crater size and photometric evolution from the Deep Impact collision. *Space Sci. Rev.* **117**, 207–239 (2005).
- Kolokolova, L., Hanner, M. S., Levasseur-Regourd, A.-Ch. & Gustafson, B. Å. S. in *Comets II* (eds Festou, M. C., Keller, H. U. & Weaver, H. A.) 577–604 (Univ. Arizona Press, Tucson, 2004).
- Fulle, M., Mikuz, H. & Bosio, S. Dust environment of comet Hyakutake 1996B2. *Astron. Astrophys.* **324**, 1197–1205 (1997).
- Fulle, M. et al. The dust tail of Comet Liller 1988 V. *Astron. Astrophys.* **253**, 615–624 (1992).
- O'Keefe, J. D. & Ahrens, T. J. Impact and explosion crater ejecta, fragment size, and velocity. *Icarus* **62**, 328–338 (1985).
- Holsapple, K. A. The scaling of impact processes in planetary sciences. *Annu. Rev. Earth Planet. Sci.* **21**, 333–373 (1993).
- Keller, H. U. in *Proc. International Workshop on Physics and Mechanics of Cometary Materials* (eds Hunt, J. & Guyeme, T. D.) 39–45 (ESA-SP 302, ESA, Publication Division, ESTEC, Noordwijk, 1989).

20. Belton, M. J. S. *et al.* Deep Impact: Working properties for the target Nucleus—Comet 9P/Tempel 1. *Space Sci. Rev.* **117**, 137–160 (2005).
21. Iseli, M., Küppers, M., Benz, W. & Bochsler, P. Sungrazing comets: Properties of nuclei and in situ detectability of cometary ions at 1 AU. *Icarus* **155**, 350–364 (2002).
22. Koschny, D. & Grün, E. Impacts into ice-silicate mixtures: Ejecta mass and size distributions. *Icarus* **154**, 402–411 (2001).

Acknowledgements OSIRIS was built by a European consortium and is part of the payload of the Rosetta spacecraft, provided by the European Space Agency. We acknowledge the work of the Rosetta Science Operations Centre and the

Rosetta Mission Operations Centre in coordinating the observation timelines and operating the spacecraft, in particular K. Wirth, V. Dhiri, P. Ferri, E. Montagnon, A. Hubault, J. Morales, V. Companys and M. Lauer. We acknowledge the funding of the national space agencies ASI, CNES, DLR, the Spanish Space Program (Ministerio de Educacion y Ciencia), SNSB and ESA.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.K. (kueppers@mps.mpg.de).

The OSIRIS Team M. F. A'Hearn¹, F. Angrilli², M. E. Bailey³, P. Barthol⁴, M. A. Barucci⁵, J.-L. Bertaux⁶, J. A. Burns⁷, G. Cremonese⁸, W. Curdt⁹, M. De Cecco¹⁰, S. Debei¹¹, M. Fulle¹², F. Gliem¹³, W.-H. Ip¹⁴, E. Kührt¹⁵, A. Llebaria¹⁶, J. J. Lopez Moreno¹⁷, F. Marzari¹⁸, G. Naletto¹⁹, L. Sabau²⁰, A. Sanz Andrés²¹, J. P. Sivan²², G. Tondello²³ & K.-P. Wenzel²⁴

Affiliations for participants: ¹University of Maryland, Department of Astronomy, College Park, Maryland 20742-2421, USA. ²Università di Padova, Dipartimento di Ingegneria Meccanica, Via Venezia, 1, 35131 Padova, Italy. ³Armagh Observatory, College Hill, Armagh BT61 9DG, UK. ⁴Max-Planck-Institut für Sonnensystemforschung, Max-Planck-Str. 2, 37191 Katlenburg-Lindau, Germany. ⁵Observatoire de Paris, Section de Meudon, 92195 Meudon, France. ⁶Services d'Aéronomie du CNRS, Route des Gatines, BP 3, 91371 Verrières le Buisson, France. ⁷Cornell University, 424 Space Sciences Building, Ithaca, New York 14853-6801, USA. ⁸Università di Padova, Dipartimento di Astronomia, Vicolo dell'Osservatorio 2, 35122 Padova, Italy. ⁹Max-Planck-Institut für Sonnensystemforschung, Max-Planck-Str. 2, 37191 Katlenburg-Lindau, Germany. ¹⁰Università di Padova, Center of Studies and Activities of Space (CISAS) "G.Colombo", Via Venezia 1, 35151 Padova, Italy. ¹¹Università di Padova, Center of Studies and Activities of Space (CISAS) "G.Colombo", Via Venezia 1, 35151 Padova, Italy. ¹²ISAS, International School for Advanced Studies, Strada Costiera, 11, 34014 Trieste, Italy. ¹³Institut für Datentechnik und Kommunikationsnetze der TU Braunschweig, Hans-Sommer-Str. 66, 38106 Braunschweig, Germany. ¹⁴National Central University, Institute of Astronomy, 32054 Chung-Li, Taiwan. ¹⁵DLR Institute for Planetary Research, Rutherfordstr. 2, 12489 Berlin, Germany. ¹⁶Laboratoire d'Astrophysique de Marseille (CNRS/LAM), Traverse du Siphon, Les Trois Lucs, BP 8, 13376 Marseille Cedex 12, France. ¹⁷Instituto de Astrofísica de Andalucía, PO Box 3004, Camino Bajo de Huétor 24, 18080 Granada, Spain. ¹⁸Università di Padova, Dipartimento di Fisica, Via Gradenigo 6A, 35151 Padova, Italy. ¹⁹Università di Padova, Dipartimento di Elettronica et Informatica, Via Venezia 1, 35151 Padova, Italy. ²⁰Instituto Nacional de Tecnica Aeroespacial, Carretera de Ajalvir, p.k. 4, 28850 Torrejon de Ardoz (Madrid), Spain. ²¹Universidad Politécnica de Madrid, Instituto Ignacio D'Riva, Plaza Cardenal Cisneros s/n, 28040 Madrid, Spain. ²²Laboratoire d'Astrophysique de Marseille (CNRS/LAM), Traverse du Siphon, Les Trois Lucs, BP 8, 13376 Marseille Cedex 12, France. ²³Università di Padova, Dipartimento di Elettronica et Informatica, Via Venezia 1, 35151 Padova, Italy. ²⁴European Space Research and Technology Centre, Space Science Department, Keplerlaan 1, Postbus 299, 2201 AZ Noordwijk ZH, The Netherlands.

Geology and insolation-driven climatic history of Amazonian north polar materials on Mars

Kenneth L. Tanaka¹

Mariner 9 and Viking spacecraft images revealed that the polar regions of Mars, like those of Earth, record the planet's climate history. However, fundamental uncertainties regarding the materials, features, ages and processes constituting the geologic record remained^{1–6}. Recently acquired Mars Orbiter Laser Altimeter data⁷ and Mars Orbiter Camera high-resolution images⁸ from the Mars Global Surveyor spacecraft and moderately high-resolution Thermal Emission Imaging System visible images⁹ from the Mars Odyssey spacecraft permit more comprehensive geologic and climatic analyses^{10–17}. Here I map and show the history of geologic materials and features in the north polar region that span the Amazonian period (~ 3.0 Gyr ago to present)^{18,19}. Erosion and redeposition of putative circumpolar mud volcano deposits¹⁵ (formed by eruption of liquefied, fine-grained material) led to the formation of an Early Amazonian polar plateau consisting of dark layered materials. Crater ejecta superposed on pedestals indicate that a thin mantle was present during most of the Amazonian, suggesting generally higher obliquity and insolation conditions at the poles than at present. Brighter polar layered deposits rest unconformably on the dark layers and formed mainly during lower obliquity over the past 4–5 Myr (ref. 20). Finally, the uppermost layers post-date the latest downturn in obliquity $<20,000$ years ago²⁰.

At the beginning of the Amazonian, outflow channel flooding and perhaps spring discharges led to thick, water-saturated sediments covering the northern plains region of Vastitas Borealis^{15,21,22} (see Fig. 1 for locations). Assuming that dark dune material in the region originates from the Vastitas Borealis material, Mars Global Surveyor Thermal Emission Spectrometer results indicate that the material may be weathered basaltic clasts²³. Otherwise, the observed²⁴ thin, recent high-latitude mantle generally obscures the spectral character and fine morphological textures of older materials in Vastitas Borealis; only their subdued morphological character can be observed. Much of Vastitas Borealis is marked by polygonal troughs indicative of compaction, and by a system of wrinkle ridges resulting from regional tectonic contraction¹⁵.

Perhaps because of this compaction and contraction, together with downward freezing of wet sediments, high subsurface fluid pressures may have developed locally. Mud volcanism may have ensued in the Scandia region, forming a few dozen pancake domes surrounded by moats tens of kilometres across and irregular complexes hundreds of kilometres across¹⁵. Terrestrial mud volcanoes form by eruption of over-pressured, water-saturated, unconsolidated sediments, and some examples have surrounding moats resulting from removal of subsurface material and surface collapse²⁵.

Extending south and east of the pancake domes, the Scandia materials include complex fields of mesas and knobs and irregular depressions indicative of highly eroded deposits associated with collapse that also may have a mud volcanic origin. Mud volcano deposits generally are fine-grained and poorly consolidated, and may

be easily eroded and transported by wind once desiccated. The eroded material may have collected at the pole and been incorporated into polar ice, as postulated¹⁶ for the origin of dark, platy basal deposits on Planum Boreum. These deposits crop out along more than half of the margin of Planum Boreum, from ~ 300 km east of Chasma Boreale westward to near the western margin of Olympia Undae. Dunes appear to emanate from several steep-sided

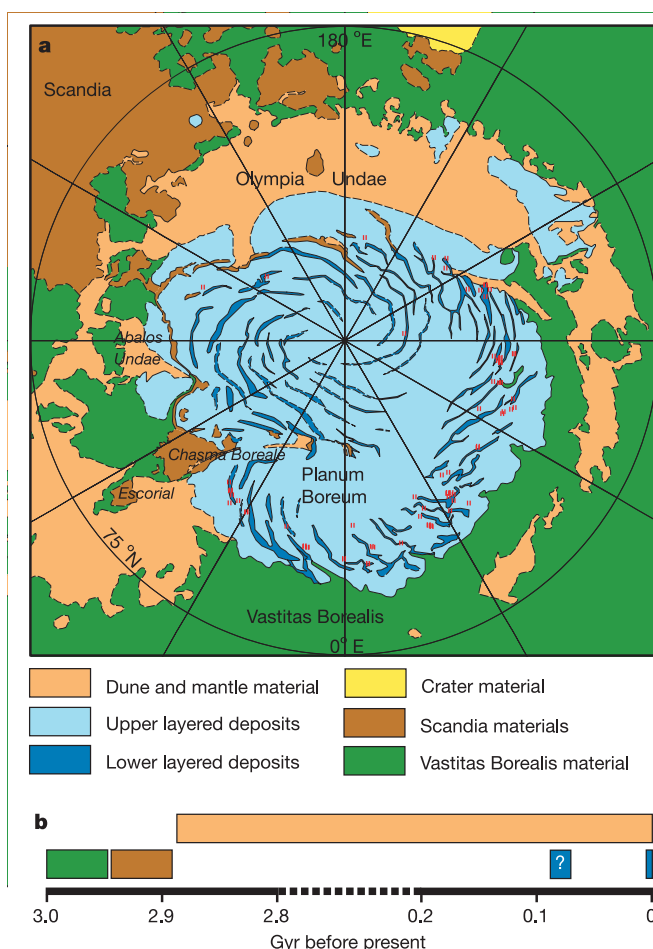


Figure 1 | Amazonian geology of the north polar region of Mars. **a**, Geologic map (red squares show locations of unconformities in MOC images; polar stereographic projection; see text for discussion of units). **b**, Timeline showing emplacement history of major north polar geologic materials (upper layered deposits are too young to be shown); older outcrops (box with question mark) similar in age to south polar layered deposits may be present.

¹Astrogeology Team, US Geological Survey, 2255 N. Gemini Drive, Flagstaff, Arizona 86001, USA.

amphitheatres hundreds of metres high along the margin of Planum Boreum (Fig. 2a), indicating a high sand content there^{8,12,16}. However, much of the deposit margin is less steep and dune-poor and may have little sand content or be mantled. Planum Boreum sands could have originated from erosion of the proposed Scandia mud vents. The largest circumpolar dune fields appear to emanate from Planum Boreum, and are partly surrounded by mantles that are sufficiently thick to subdue the polygonal troughs in the underlying surface. It may be that these mantles consist mostly of fines generated by comminution of dune sands rather than by fallout from the atmosphere².

Some of the Scandia materials could be finer-grained loess that could have travelled great distances by wind from source regions. Candidate loess deposits include possible sand-poor parts of the Scandia materials that form locally more than 1,000 m of the base of Planum Boreum. These materials are best exposed in the 200–350-m-thick plateau that forms the floor of Chasma Boreale. The plateau is made up of ~10 layers visible in Thermal Emission Imaging System (THEMIS) images (Fig. 2b), and is marked by the same suite of polygonal troughs that deform the adjacent Vastitas Borealis surface. South of Chasma Boreale, the 22-km-diameter Escorial crater (77°N, 305°E) and its ejecta apparently armour an isolated, 700-m-thick mesa¹⁶. Because the Scandia materials are the only plateau-forming materials in the area, the Escorial mesa is probably a remnant of that unit. From Escorial down to a complex of linear scarps at 67–69°N that form an apparent erosional margin of the unit, Scandia materials appear to be almost completely removed.

The average depth of erosion of Scandia materials, which cover a

total of $\sim 1.5 \times 10^6 \text{ km}^2$ in Vastitas Borealis, is $>100 \text{ m}$, based on the heights of the knobs and mesas that range from tens to hundreds of metres. This yields a displaced volume of $\sim 1.5 \times 10^5 \text{ km}^3$. The Scandia materials of Planum Boreum probably include a lower sequence of primary deposits (as in the floor of Chasma Boreale) overlain by material redeposited from eroded Scandia materials in Vastitas Borealis. Given the areal extent of the reworked material of $\sim 5 \times 10^5 \text{ km}^2$ and a mean thickness of 300 m, the volume of reworked material approximately matches the estimated eroded volume of Scandia materials from Vastitas Borealis. Although extensive resurfacing of Scandia materials precludes derivation of precise relative ages based on crater densities, Scandia materials both in Vastitas Borealis and on Planum Boreum include superposed impact craters $>10 \text{ km}$ in diameter (Fig. 2c), indicating that the materials probably date back to the Early Amazonian, perhaps not long after emplacement of underlying plains sediments.

Most craters in Vastitas Borealis surrounding Planum Boreum and within a depression at 84.8°N, 2.5°E atop the eroded Scandia materials at the head of Chasma Boreale occur on pedestals mostly $<30 \text{ m}$ high (Fig. 2d) and rarely 50–100 m high. The pedestal craters show that the Scandia and Vastitas Borealis surfaces were thinly mantled during most of the Amazonian. No large pedestal or exhumed craters have been found that clearly superpose the thick bright, polar layered deposits (PLD), but thin PLD in a crater pedestal may be indistinguishable from a thin mantle. Thus, thick Amazonian PLD generally have been short-lived and/or restricted to about their present distribution¹⁶. These findings suggest that polar dust deposits were widely dispersed during most of the Amazonian,

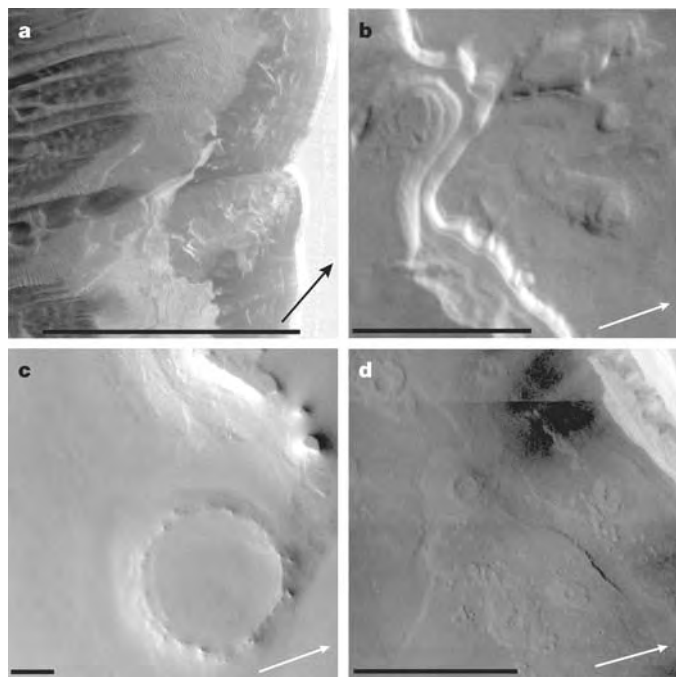


Figure 2 | Early Amazonian north polar features on Mars. Scale bars at the bottom of each image are 2 km; north is at the top; arrows show illumination direction. **a**, A source area for the Abalos Undae dune field, which results from erosion of platy Scandia materials at the base of the scarp. Mosaic of portions of MOC images R01-00086 and R01-00728; 81.8°N, 282.6°E. **b**, A lower section of Scandia materials showing ten distinct layers. Portion of THEMIS visual image V10538002; 80.8°N, 301.5°E. **c**, An impact crater 6.6 km in diameter on Planum Boreum. The crater is superposed on Scandia materials, which form mesas at upper right. Layered deposits bury and embay the Scandia materials and the crater rim. Portion of THEMIS visual image V10551038; 82.4°N, 290.4°E. **d**, The floor of a depression at the head of Chasma Boreale shows pedestal craters on top of lower Scandia materials. Portion of THEMIS visual image V04570011; 84.8°N, 3.0°E.

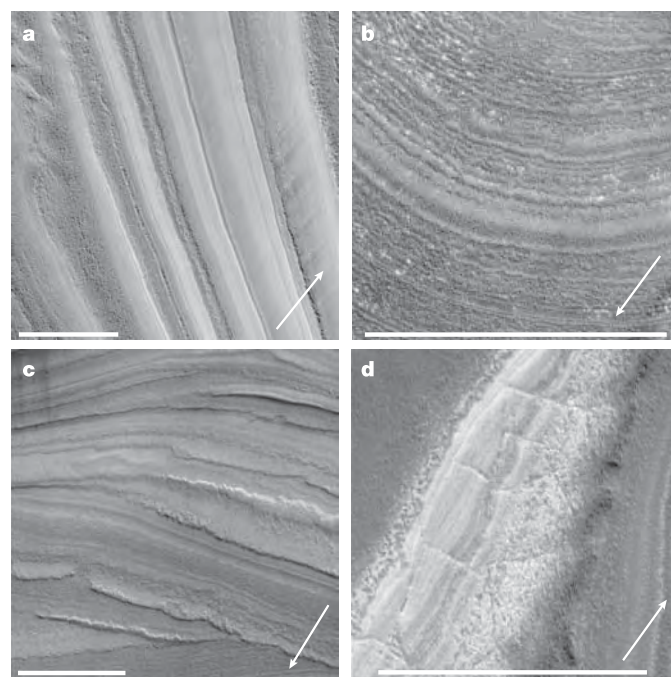


Figure 3 | Geologic characteristics of north polar lower layered deposits (LLD) of Planum Boreum. Scale bars at the bottom of each image are 1 km; north is towards the top; arrows show illumination direction. **a**, A series of seven terraces in LLD defined by brightly lit, sun-facing scarps and darker, gently sloping surfaces. Additionally, surface brightness is affected by texture, layering, frost and streaks of surficial fines. Portion of MOC image E03-00301; 82.7°N, 48.7°E. **b**, Fine-scale layering in LLD displayed on a 3–7° slope. Each layer is ~ 10 – 90 m across and ~ 0.5 – 10 m thick. Portion of MOC image R01-00384; 85.1°N, 193.3°E. **c**, A series of unconformities in LLD. Two layers show possible fractures and offsets. Portion of MOC image M22-02387; 82.4°N, 72.8°E. **d**, This complexly faulted, bright sequence of LLD may be the same fractured sequence as in **c**. The faulted sequence is unconformably overlain and embayed by younger, darker LLD. Portion of MOC image R01-00261; 82.4°N, 74.5°E.

which may be explained by higher obliquity that increases high-latitude insolation and decreases latitudinal variation in total insolation²⁶. In such conditions, polar surface ice probably becomes unstable and seasonal frost becomes longer-lived across a broad latitude range, where it may capture and deposit dust.

The PLD bury the Scandia materials of Planum Boreum and form several scattered outliers in surrounding plains down to 73° N. Previous geologic mapping separated the bright PLD from the residual ice⁵. Here, I group the residual ice with the uppermost PLD layers as the upper layered deposits (ULD). The residual ice commonly has a gradational, complex margin^{3,10} and thus is difficult to map separately from layered deposits. The close spatial correlation of the residual ice and underlying layers supports the hypothesis that the dust was fixed to the surface by ice². The ULD, which cover most of Planum Boreum, consist of one to several layers observed in Mars Orbiter Camera (MOC) images of defrosted surfaces. ULD occur on pole-facing scarps and trough floors, where the unit lies unconformably on underlying lower layered deposits (LLD; for example, MOC image R01-00556). The LLD are exposed mainly in equator-facing scarps that bound the spiral to irregular troughs and margins of Planum Boreum. In MOC images, LLD brightness varies according to the albedo of layers and surficial deposits (including wind streaks and frost), texture and illumination (Fig. 3). The upper 800 m of the north PLD can be divided into four zones of varying brightness character versus depth¹⁷, with the uppermost 300-m-thick zone bearing a dominant ~30-m wavelength signature that may represent a 51-kyr insolation cycle. Terraces also discriminate layer sequences; several to a dozen or more terraces can be observed in many Mars Orbiter Laser Altimeter slope profiles of the LLD (Fig. 3a). More than 250 individual layers mostly 1–10 m thick occur in better exposures (Fig. 3b). The correlation of terraces and individual layers with albedo and climate variations is undetermined.

Unconformities within the LLD⁸ demonstrate discontinuous patterns of layer deposition. In scarps and troughs of the LLD, 75 unconformities have been identified thus far in MOC images (Fig. 1). Most unconformities occur in lower sections of LLD and in outer troughs in the half of Planum Boreum east of Chasma Boreale, where Scandia materials are thin or absent. Possibly, areas of thicker Scandia materials provided a raised platform upon which LLD were less exposed to saltating sand, whereas LLD on the Vastitas Borealis surface and on thinner Scandia materials were subjected to much more sand abrasion. The orientations of unconformity planes and overlying LLD seem to occur at random, and preferred accumulation of LLD on pole-facing scarps is not evident. In one case, a series of unconformities occur on the southwest side of a shallow saddle on a ridge between two parallel troughs, where the northeastern trough ends (Fig. 3c). Apparently, periods of relatively enhanced erosion and/or lack of deposition of LLD at the saddle account for the unconformities, which could have been caused by winds that funnelled through the troughs and the saddle. Wind funneling may also account for troughs that cut entirely through the LLD and expose trough-aligned scarps within underlying Scandia materials (as at 83.8° N, 115.8° E and adjacent troughs to the west). These outcrops constrain the possible retreat of equator-facing LLD across the flat exposed surfaces of the Scandia material to <4 km. These observations, along with the lack of significant, preferred warming of equator-facing trough walls based on THEMIS-derived surface temperatures²⁷, indicate that insolation-driven trough migration^{1,3} has been ineffective, however.

Clear examples of structural deformation in available data are rare in the LLD. The only unambiguous structures I have observed in MOC images are fault systems displaying offsets of tens of metres within two outcrops of a relatively bright, lower LLD sequence that are each 12 km from the crest of a 43-km-diameter impact crater (Fig. 3c, d). The coincidence of the faulting and the crater suggests that deformation associated with the crater, either during or after impact, produced the faults. The sequence is unconformably overlain

by unfaulted LLD and may be a remnant of a much older LLD sequence.

Only two craters >330 m in diameter have been discovered on the LLD in MOC images M20-00372 (ref. 13) and R01-00641 acquired through relay sub-phase R02, which cover about $2.73 \times 10^4 \text{ km}^2$ of the $8.52 \times 10^4 \text{ km}^2$ exposed LLD. Using a -2 power-law crater production function and a martian cratering rate equivalent to twice that of the Moon⁶ yields a mean surface age of 3.6 ± 2.5 Myr. Given that the reliability of the LLD cratering record is uncertain, the LLD may correspond to the past 4–5 Myr period of low insolation¹⁴, forming by transfer of dust from the broad, thin high-latitude mantle evidenced by pedestal craters and from other lower latitude dust sources. Higher insolation characterizes the period from 5 to 20 Myr ago, before which computed solutions of insolation based on orbital and obliquity behaviour become chaotic²⁰. However, accumulation of much of the south PLD estimated at >30–100 Myr ago (ref. 13) suggests lower polar insolation and possible simultaneous emplacement of north PLD. Some of the sequences of north PLD lying below unconformities may be of this age (Fig. 1b).

Two craters 42 and 90 m in diameter in MOC images M19-01628 and M21-00745, respectively, on the $1.77 \times 10^5 \text{ km}^2$ of ULD (total area, $8.89 \times 10^5 \text{ km}^2$) covered by MOC images through relay sub-phase R02 indicate that the unit has a mean age of $\sim 8,700 \pm 6,200$ yr. Thus ULD have been forming during the latest downtrend in insolation that began $\sim 20,000$ yr ago¹⁴. Furthermore, each resolved ULD layer would represent a mean period of deposition of the order of 10^3 yr and have a mean deposition rate of $\sim 1\text{--}10 \text{ mm yr}^{-1}$. The origin of individual layers is uncertain, and they are thick enough to require accumulation from many dust storms. Interestingly, a 1,470-yr climate cycle of unknown origin is evident on Earth, where Greenland ice cores and other sources reveal warming events defined by peaks in $\delta^{18}\text{O}$ abundance²⁸. If the same cycle controls the thinnest layers of the LLD, then an external forcing mechanism that can affect both planets is required, such as variations in solar luminosity²⁸.

Received 1 March; accepted 19 July 2005.

- Howard, A. D. Origin of the stepped topography of the Martian poles. *Icarus* **34**, 581–599 (1978).
- Squyres, S. W. The evolution of dust deposits in the Martian north polar region. *Icarus* **40**, 244–261 (1979).
- Howard, A. D., Cutts, J. A. & Blasius, K. R. Stratigraphic relationships within Martian polar-cap deposits. *Icarus* **50**, 161–215 (1982).
- Cutts, J. A. & Lewis, B. H. Models of climatic cycles recorded in Martian polar layered deposits. *Icarus* **50**, 216–244 (1982).
- Tanaka, K. L. & Scott, D. H. *Geologic Map of the Polar Regions of Mars* (Misc. Invest. Ser. Map I-1802-C, US Geol. Surv., Reston, Virginia, 1987).
- Herkenhoff, K. E. & Plaut, J. J. Surface ages and resurfacing rates of the polar layered deposits on Mars. *Icarus* **144**, 243–253 (2000).
- Smith, D. E. et al. Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. *J. Geophys. Res.* **106**, 23689–23722 (2001).
- Malin, M. C. & Edgett, K. S. Mars Global Surveyor Mars Orbiter Camera: Interplanetary cruise through primary mission. *J. Geophys. Res.* **106**, 23429–23570 (2001).
- Christensen, P. R. et al. Morphology and composition of the surface of Mars: Mars Odyssey THEMIS results. *Science* **300**, 2056–2061 (2003).
- Thomas, P. C. et al. North-south geological differences between the residual polar caps on Mars. *Nature* **404**, 161–164 (2000).
- Kolb, E. J. & Tanaka, K. L. Geologic history of the polar regions of Mars based on Mars Global Surveyor data: II. Amazonian Period. *Icarus* **154**, 22–39 (2001).
- Byrne, S. & Murray, B. C. North polar stratigraphy and the paleo-erg of Mars. *J. Geophys. Res.* **107**, doi:10.1029/2001JE001615 (2002).
- Koutnik, M., Byrne, S. & Murray, B. South polar layered deposits of Mars: The cratering record. *J. Geophys. Res.* **107**, doi:10.1029/2001JE001805 (2002).
- Laskar, J., Levrard, B. & Mustard, J. F. Orbital forcing of the martian polar layered deposits. *Nature* **419**, 375–377 (2002).
- Tanaka, K. L., Skinner, J. A. Jr, Hare, T. M., Joyal, T. & Wenker, A. Resurfacing history of the northern plains of Mars based on geologic mapping of Mars Global Surveyor data. *J. Geophys. Res.* **108**, doi:10.1029/2002JE001908 (2003).
- Fishbaugh, K. E. & Head, J. W. III Origin and characteristics of the Mars north polar basal unit and implications for polar geologic history. *Icarus* **174**, 444–474 (2005).

17. Milkovich, S. M. & Head, J. W. III North polar cap of Mars: Polar layered deposit characterization and identification of a fundamental climate signal. *J. Geophys. Res.* **110**, E01005, doi:10.1029/2004JE002349 (2005).
18. Tanaka, K. L. The stratigraphy of Mars. *J. Geophys. Res.* **91**, E139–E158 (1986).
19. Hartmann, W. K. & Neukum, G. Cratering chronology and the evolution of Mars. *Space Sci. Rev.* **96**, 165–194 (2001).
20. Laskar, J. *et al.* Long term evolution and chaotic diffusion of the insolation quantities of Mars. *Icarus* **170**, 343–364 (2004).
21. Parker, T. J., Saunders, R. S. & Schneeberger, D. M. Transitional morphology in west Deuteronilus Mensae, Mars: Implications for modification of the lowland/upland boundary. *Icarus* **82**, 111–145 (1989).
22. Baker, V. R. *et al.* Ancient oceans, ice sheets, and the hydrological cycle on Mars. *Nature* **352**, 589–594 (1991).
23. Wyatt, M. B., McSween, H. Y. Jr, Tanaka, K. L. & Head, J. W. III. Global geologic context for rock types and surface alteration on Mars. *Geology* **32**, 645–648 (2004).
24. Mustard, J. F., Cooper, C. D. & Rifkin, M. K. Evidence for recent climate change on Mars from the identification of youthful near-surface ground ice. *Nature* **412**, 411–414 (2001).
25. Robertson, A. & Ocean Drilling Program Leg 160 Scientific Party, Mud volcanism on the Mediterranean Ridge: Initial results of Ocean Drilling Program Leg 160. *Geology* **24**, 239–242 (1996).
26. Mellon, M. T. & Jakosky, B. M. The distribution and behaviour of Martian ground ice during past and present epochs. *J. Geophys. Res.* **100**, 11781–11799 (1995).
27. Fishbaugh, K. E. Characterization of Martian north polar geologic units using Mars Odyssey THEMIS data. *Lunar Planet Sci. Conf.* [CD-ROM] XXXVI, Abstr. no. 1335 (2005).
28. Rahmstorf, S. Timing of abrupt climate change: A precise clock. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017115 (2003).

Acknowledgements I acknowledge discussions and other assistance from S. Byrne, K. Fishbaugh, E. Kolb, K. Herkenhoff, A. Rodriguez, J. Skinner and T. Hare. This work was supported by NASA.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing financial interests. Correspondence and requests for materials should be addressed to the author (ktanaka@usgs.gov).

Coherent signal amplification in bistable nanomechanical oscillators by stochastic resonance

Robert L. Badzey¹ & Pritiraj Mohanty¹

Stochastic resonance^{1,2} is a counterintuitive concept: the addition of noise to a noisy system induces coherent amplification of its response. First suggested as a mechanism for the cyclic recurrence of ice ages, stochastic resonance has been seen in a wide variety of macroscopic physical systems: bistable ring lasers³, superconducting quantum interference devices^{4,5} (SQUIDs), magnetoelastic ribbons⁶ and neurophysiological systems such as the receptors in crickets⁷ and crayfish⁸. Although fundamentally important as a mechanism of coherent signal amplification, stochastic resonance has yet to be observed in nanoscale systems. Here we report the observation of stochastic resonance in bistable nanomechanical silicon oscillators. Our nanomechanical systems consist of beams that are clamped at each end and driven into transverse oscillation with the use of a radiofrequency source. Modulation of the source induces controllable switching of the beams between two stable, distinct states. We observe that the addition of white noise causes a marked amplification of the signal strength. Stochastic resonance in nanomechanical systems could have a function in the realization of controllable high-speed nanomechanical memory cells, and paves the way for exploring macroscopic quantum coherence and tunnelling.

The conditions for a system to exhibit stochastic resonance are well known and fairly robust. First among these is an energy threshold. Second, there needs to be an applied subthreshold modulation, which causes the system to cross the threshold. And last, there needs to be a source of noise. One of the classic examples of a system that undergoes stochastic resonance is that of a particle in a double-well potential. With no modulation, the particle will undergo a transition between the two wells with the Kramers rate, $\Gamma_K \propto e^{-\Delta V/D}$, where ΔV is the depth of each well and D is the strength of temperature-induced noise. With the addition of external periodic modulation there is an interaction between the modulation and the noise, resulting in a resonance in the signal-to-noise ratio (SNR), which is defined² as

$$\text{SNR} = 2 \left[\lim_{\Delta\omega \rightarrow 0} \int_{\Omega - \Delta\omega}^{\Omega + \Delta\omega} S(\omega) d\omega \right] / N(\Omega) = \frac{S(\Omega)}{N(\Omega)} \quad (1)$$

where $S(\Omega)$ is the height of the power spectrum peak at the modulation frequency Ω and $N(\Omega)$ is the spectral background.

It is well known that, with suitably strong driving, the linear response of a damped, driven harmonic oscillator undergoes a transition to that of a nonlinear bistable system, which is the famous Euler instability. The equation of motion is described by the Duffing equation with driving force F and driving amplitude ω :

$$\ddot{x} + \gamma\dot{x} + kx \pm k_3x^3 = F \cos \omega t \quad (2)$$

Here, k and k_3 are the linear and nonlinear spring constants, respectively, and γ is the damping coefficient describing the dissipation of the oscillator. The sign of the k_3 term changes depending on

whether the oscillator is under compressive (positive k_3) or tensile (negative k_3) strain. The presence of the cubic term in the equation of motion naturally leads to a term of fourth order in the beam potential, which implies that the oscillator now demonstrates bistability^{9–11}. It is important to note that, regardless of the sign of the nonlinear term, both instances will result in a bistable oscillator. This is clear from the amplitude response of the Duffing oscillator: when looking at the amplitude response as a function of frequency, one finds that the oscillator no longer displays the symmetric lorentzian line shape characteristic of a damped, driven, linear harmonic oscillator. Rather, there is a sharp bend in the peak that creates a region of frequency space in which the oscillator is multi-valued. If the oscillator is softening, the bend occurs to the left of the location of the linear peak; similarly, the bend is on the right for a hardening spring. The signature of a system in nonlinear response is a hysteresis defined by sweeping the driving frequency forwards and backwards through the bistable region. For instance, in the compressive case, as the frequency is increased, the response will follow state 1 until reaching a point of instability, whereupon it will drop sharply to state 2. A downward sweep will produce the opposite effect, defining a region in which the beam is bistable; these situations are reversed for a softening spring. Excitation at a single frequency within this region allows one to access either of the two bistable states by the addition of a suitable modulation¹². In each case, bistable behaviour holds only for driving frequencies near resonance.

Strictly speaking, the above description is that of a point particle in a double-well potential with no additional modulation or noise terms. For a spatially extended system such as a doubly clamped beam or string under additional modulation and external noise, the more appropriate description is the Landau–Ginzburg equation, in which the field variable $\Phi(x)$ replaces the position variable x :

$$\frac{d\Phi(x)}{dt} = m\Phi(x) - \Phi^3(x) + \kappa \frac{d^2\Phi(x)}{dx^2} + F_{\text{mod}} \cos \Omega t + \eta(x, t) \quad (3)$$

Here Ω is the modulation frequency and $\eta(x, t)$ is the added noise term. In their seminal paper, Benzi *et al.*¹³ predicted the emergence of stochastic resonance in a doubly clamped beam or string described by the Landau–Ginzburg equation with Von Neumann boundary conditions, $d\Phi(0)/dx = d\Phi(L)/dx = 0$. In addition to the two stable homogeneous solutions corresponding to the occupation of either well, there also exists a class of instanton solutions which imply that in the presence of noise the string can undergo a transition between the two wells. In this case, the string collective coordinate $u = (1/L) \int_0^L \Phi(x) dx$ synchronizes with the modulation within a certain range of noise powers. The behaviour of the collective coordinate as a particle in a double-well potential therefore completes the projection of the doubly clamped beam into the canonical single-particle picture described above.

We have fabricated two doubly clamped nanomechanical beams from single-crystal silicon with e-beam lithography and dry etching.

¹Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA.

Frequency sweeps with increasing drive amplitude show a marked evolution from the familiar lorentzian response into a shifted response with a sharp drop and marked hysteresis. We perform forward and reverse frequency sweeps with large driving amplitudes to confirm and record the onset of nonlinear and bistable behaviour. We then drive each bridge at a single frequency within its unique bistable frequency region to access both bistable states by means of the addition of a modulation signal of suitable strength. Our samples are driven with the well-known magnetomotive technique¹⁴; Fig. 1 shows a diagram of the circuit used to drive the bistable oscillator and add white noise.

Beam 1 is cooled to 300 mK and driven nonlinear with $A_{\text{drive}} = 4.0$ dBm. The modulation is made too weak (-5.5 dBm) to induce switching, and the noise is swept from -71 dBm (~ 79 pW) to -41 dBm (~ 79 nW). At each noise power, a time series is taken to monitor the oscillator response. Figure 2A shows selected time series at different noise powers, with their associated power spectra. After calculating the SNR, the data are combined and represented in Fig. 2B. The resonance is clear and marked, with a sharp increase in SNR at ~ 18 nW. These noise powers are those coming from the source rather than the actual power incident on the sample.

It has been shown¹⁵ that the temperature of the beam is a viable source of noise, causing a decrease in switching fidelity and increase in state noise at temperatures below 1 K. There should therefore be a temperature or range of temperatures above 1 K where we would actually see an increase in the SNR. Beam 2 is disconnected from the noise source. In this case, the beam achieves a nonlinear response with a smaller driving amplitude (1.0 dBm), whereas the amplitude necessary for switching is much greater (>16.5 dBm). Because the modulation power is so large, the electrical signal actually shows up in our lock-in technique, as can be seen in Fig. 3A. Again, the left-hand side shows segments from the time series at each temperature

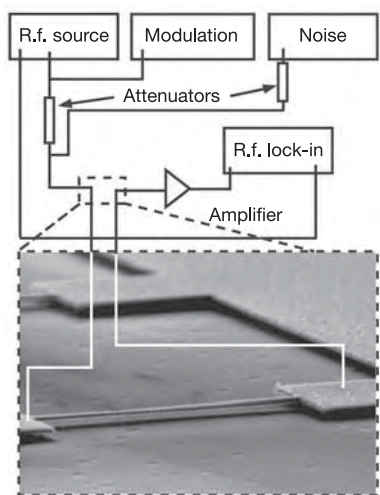


Figure 1 | Diagram of measurement circuit. The circuit serves to excite the nonlinear motion of the beam, modulate the switching, and add in noise to achieve stochastic resonance. The radiofrequency (r.f.) source is a Rohde & Schwarz function generator, operating at a frequency within the bistable region (either 23.4973 or 20.8343 MHz, depending on the beam), at a driving amplitude of 4.0 dBm (beam 1) or 1.0 dBm (beam 2). The modulation is produced by an HP 3325 synthesizer, with an amplitude of -5.5 dBm (beam 1) or 16.5 dBm (beam 2). In both cases the modulation is a square wave with a frequency of 0.05 Hz. The white-noise source is an HP 33220 synthesizer, capable of creating white noise in a band between 0 and 15 MHz. For the temperature sweep (beam 2) the noise source was disconnected. The attenuator is an 8 k Ω resistor; the noise source has a separate 30 dB attenuator at the source output. Our beams are 7–8 $\mu\text{m} \times 300 \text{ nm} \times 200 \text{ nm}$ in dimension, and are placed in a ^3He cryostat (300 mK to 4 K) at the centre of a 9-T superconducting solenoid magnet.

(with the modulation amplitude highlighted in grey), whereas the right-hand side shows the power spectra obtained by fast Fourier transform (FFT). It is immediately apparent that the switching is much noisier than that of beam 1, even when one takes into account the background effect of the modulation. Full switching is never really achieved, even for the most responsive time series.

The difficulty in inducing switching behaviour can be understood by considering the symmetry of the double-well potential. The response of beam 1 led to the conclusion that the potential is very nearly symmetric. However, it rapidly became apparent that this was not true of beam 2. This is especially evident when looking at

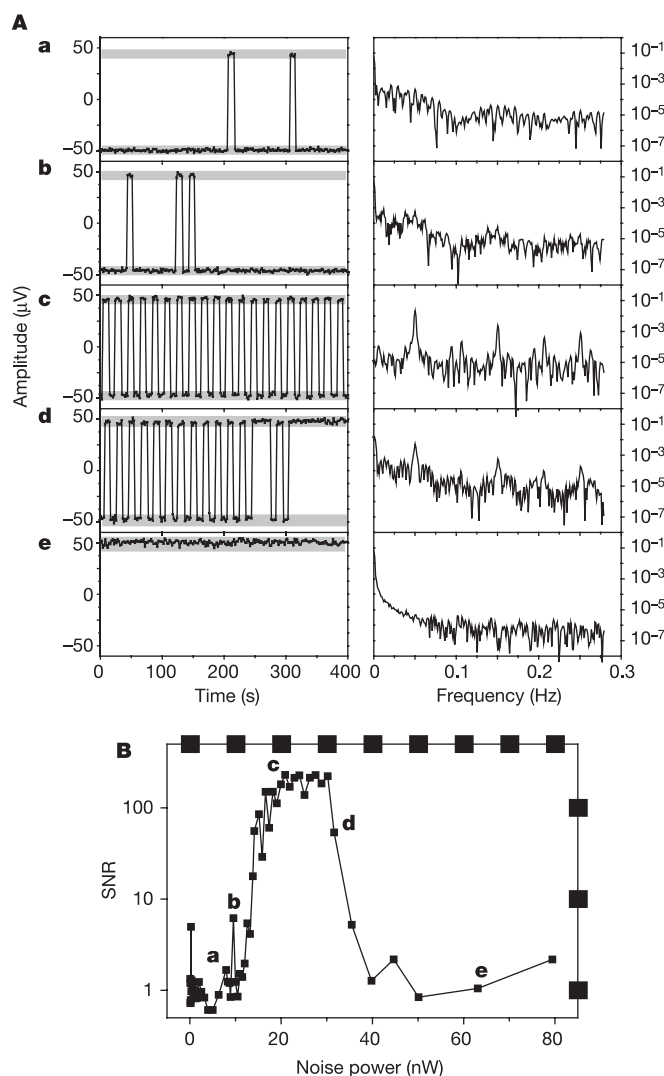


Figure 2 | Re-emergence of switching behaviour as a function of added white noise on beam 1 ($f_{\text{drive}} = 23.4973$ MHz). **A**, The left-hand graphs show selected beam responses as a function of added noise power, with lower powers at the top. The grey bands represent the noise of each state, including contributions to the electrical signal from the modulation itself. There is no response until ~ 10 nW of power is introduced (**a**), when a few sporadic switches are seen. As expected, the power spectrum of the response shows no dominant spectral components. With increasing noise power, we observe an increase of switch events, as shown in **b** (13 nW) and **c** (27 nW). Finally, as the noise is increased still further, the switches begin to die out (**d**, 32 nW) until a regime of no switching occurs (**e**, 63 nW). The right-hand column of plots shows the FFT of the data in the left column. As can be expected, there are sharp peaks at 0.05 Hz and all odd harmonics thereof, as befits the square-wave response of the switch. **B**, Each individual scan was processed to obtain the SNR for each noise power. The letters **a–e** correspond to the scans in **A**.

Fig. 3A,b. At each impulse from the modulation there is a sharp spike up to the upper state, which immediately decays down to the lower state; this is an indication that the potential of the bridge is asymmetric. The occurrence of incoherent switches down to the lower state is so prevalent that even when the SNR is maximized, the switching is not truly coherent. This is in sharp contrast to the very clean, symmetric behaviour shown by beam 1. In addition, a look at the power spectra from bridge 2 reveals the presence of even-order harmonics, a sure signature of an asymmetric potential^{16,17}. The spring coefficients k_1 and k_3 change with temperature, which might affect the symmetry and even the shape of the potential. Such a change is manifested by a shift of the resonance frequency of the

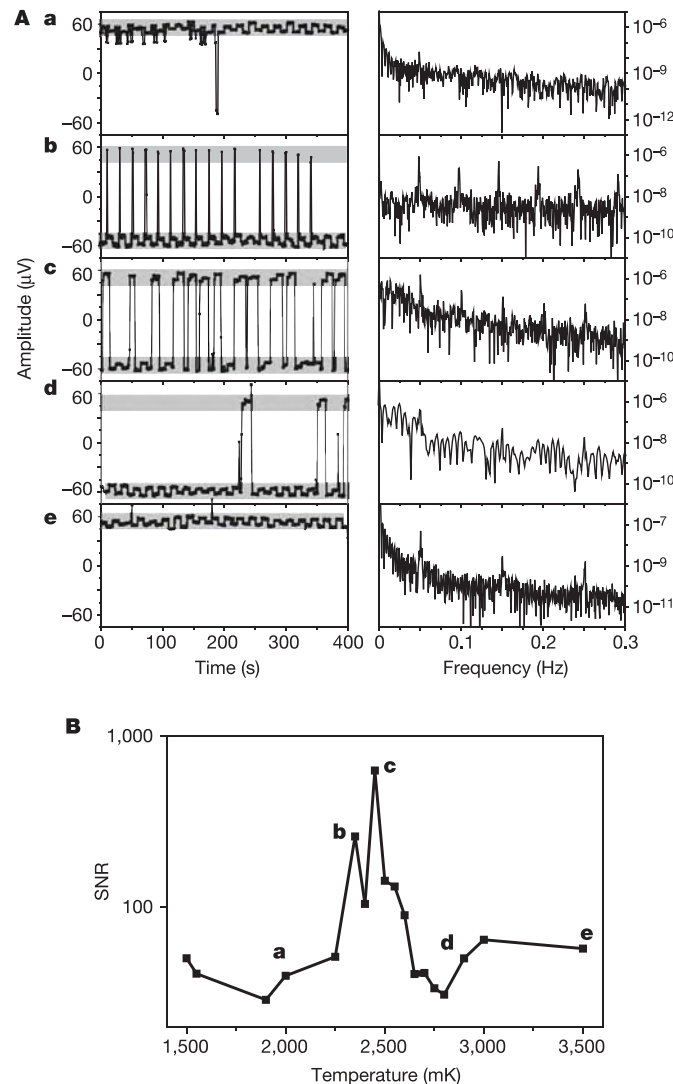


Figure 3 | Switching behaviour as a function of temperature on beam 2 ($f_{\text{drive}} = 20.8348 \text{ MHz}$). **A**, As the temperature is increased, switching behaviour again re-emerges. The modulation is seen in all graphs. Again, the right-hand column shows the FFT of the left-hand column, demonstrating the coherence of the signal as the switches re-emerge. The frequency spectrum shows many more peaks than one would naively expect from a simple square-wave response. This is at least partly due to the increased number of incoherent switches evidenced in this beam. **a**, 2,000 mK; **b**, 2,450 mK; **c**, 2,600 mK; **d**, 2,800 mK; **e**, 3,500 mK. **B**, The SNR was obtained from each individual scan, with a clear resonance occurring at $\sim 2,600 \text{ mK}$ (2.6 K). The line simply connects each individual temperature point, with no fit involved or implied. In contrast to the very broad resonance seen with noise power, this peak is quite sharp and dies off very quickly. This is not unexpected, because the sweep in temperature is much coarser than that used with noise.

beam with a change in temperature, which is proportional to the sound velocity $v = \sqrt{E/\rho}$, where E is the Young modulus and ρ is the material density. As the temperature is reduced, the modulus is increased, leading to a subsequent increase in the resonance frequency. However, we found that such frequency shifts were on the order of less than 10^{-4} within the temperature range we employed during these experiments.

Another important question is the effect of temperature on the dissipation coefficient γ . Should the effect be non-negligible, there exists a multiplicative noise term in the Duffing equation. This is in sharp contrast to the purely additive noise of the classic theory of stochastic resonance. The dissipation is most easily quantified by measuring the quality factor Q of the oscillator as a function of temperature. Our measurements of this quantity indicate that Q can change by as much as 10% over the temperature range 2–3 K.

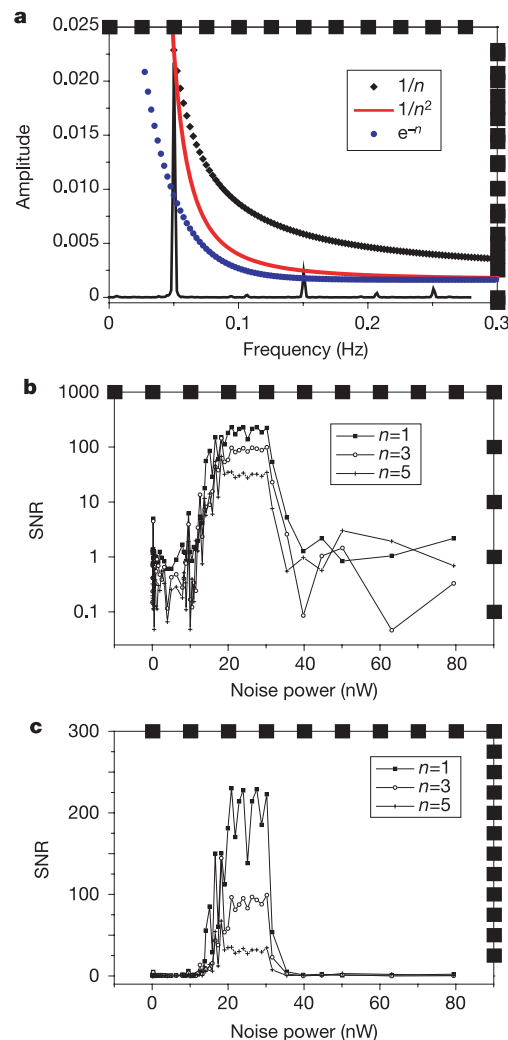


Figure 4 | Behaviour of higher harmonics. **a**, The power spectrum from a representative time series, showing the presence of spectral peaks at the odd harmonics of the modulation frequency. This is one of the hallmarks of the square wave drive and response; a sine wave would have a single fundamental peak at the modulation frequency. The three curves represent fits of the peak height versus n , the harmonic number: black diamonds, $1/n$; red line, $1/n^2$; blue circles, e^{-n} . As expected, the peak height varies as $1/n^2$. **b**, Comparison of the SNR for the higher harmonics with the fundamental. Filled squares, $n = 1$; open circles, $n = 3$; crosses, $n = 5$. For comparison, all three graphs are plotted at the same scale as Fig. 3b. **c**, The SNR for each frequency is plotted again, with the y axis plotted on a linear, rather than a logarithmic, scale, to emphasize the difference in SNR for each frequency. Symbols as in **b**.

Although this is not a huge effect, it can certainly have an impact on the system noise, and is definitely worth exploring in more detail. The topic of stochastic resonance has been broached in the literature^{18,19}, but this provides an avenue for continued research (A. Bulsara, personal communication).

Again, an analysis of each power spectrum produces the SNR value, and these points are compiled to create Fig. 3B. The peak of SNR is much sharper than that seen in beam 1. This might at least in part be due to the rather coarse sweep, when looked at from the perspective of input noise. The apparent increase in SNR at higher temperatures (>3,000 mK) is somewhat misleading. This is due to the still-present signal from the modulation, which produces a peak in the power spectrum and is not due to actual switching between the two bistable states. Filtering the data to excise the effect of modulation would undoubtedly produce a cleaner power spectrum, but the data are presented raw to underscore the presence of stochastic resonance in a system that is less than optimal in potential symmetry and background noise.

In sharp contrast to most other systems investigated for stochastic resonance, we use a square-wave modulation, as opposed to the canonical sine wave. This is an experimental consideration, because we found that switching was much more easily achieved with a square wave. Although a seemingly minor technical difference, there are some important ramifications of the use of this type of modulation. Among these, the modulation signal can be decomposed into the form

$$S_{\text{mod}} = \sum_{n=1}^{\infty} a_n \cos n\Omega t + b_n \sin n\Omega t \quad (4)$$

Here S_{mod} is the input signal, n is the harmonic number, and Ω is the modulation angular frequency. For a linear response system that follows ref. 20 this does not significantly alter the SNR behaviour; the ratio is still dominated by the fundamental frequency. However, in the presence of a square wave there should be amplifications at other harmonics (M. Grifoni, personal communication) of the fundamental and they should be related to the fundamental SNR as $\text{SNR}(\Omega)/\text{SNR}(n\Omega) = n^2$. For a square wave, only the odd harmonics should be present. Figure 4a shows a representative power spectrum from beam 1, plotted on a linear scale to emphasize the peak height. As can be seen, the peak heights scale most closely as $1/n^2$. For a flat background noise spectrum, peak height and SNR are the same quantity. In addition, Fig. 4b, c shows the SNRs from the fundamental and the first two odd harmonics. Figure 4b shows all three on the same scale as Fig. 2B, whereas Fig. 4c shows them plotted on a linear scale to emphasize the difference in the SNR. Theoretical investigations in this area are few^{21,22} but we hope that in the light of this experiment more work might be done making this distinction.

It is typical that after exceeding the maximum noise needed to induce stochastic resonance, the time scan would be characterized by a very noisy response dominated by incoherent switches caused by the large noise power. In contrast, both of these systems show no switching at all when reaching the upper limit of noise power, instead settling on one state. As mentioned above, this is not a system with a static bistability—rather, its bistable and nonlinear behaviour is induced by the presence of a strong drive. This drive and the hysteresis it creates impart a certain amount of rigidity to the system, making switching between the two states difficult. In comparison with the size of the signal required to both form the bistable two-state system and to achieve full switching in the absence of noise (~1–10 dBm), the maximum amount of noise induced either externally or by temperature is orders of magnitude smaller. This is certainly a topic worthy of future investigation, because this dynamic bistability is one of the fundamental aspects that departs from the canonical stochastic resonance model.

The presence of stochastic resonance in these systems is exciting

not only for its inherent physical interest but also because it brings up the possibility of using stochastic resonance as a means of enhancing signal processing. If the nanomechanical memory element does indeed turn out to be a viable technological innovation, stochastic resonance may very well be one of the mechanisms that permit its everyday use. Even though our system is manifestly classical, it is closely related to at least one quantum-mechanical system (the macroscopic quantum harmonic oscillator²³). The demonstration of stochastic resonance in this system enables the exploration of stochastic resonance in a quantum-mechanical context on a system that is at the very forefront of quantum logic, quantum control and quantum computation.

Received 2 June; accepted 11 August 2005.

1. Benzi, R., Sutera, A., Parisi, G. & Vulpiani, A. A theory of stochastic resonance in climatic change. *J. Appl. Math.* **43**, 565–578 (1983).
2. Gammaitoni, L., Hänggi, P., Jung, P. & Marchesoni, F. Stochastic resonance. *Rev. Mod. Phys.* **70**, 223–287 (1998).
3. McNamara, B., Wiesenfeld, K. & Roy, R. Observation of stochastic resonance in a ring laser. *Phys. Rev. Lett.* **60**, 2626–2629 (1988).
4. Hibbs, A. D. *et al.* Stochastic resonance in a superconducting loop with a Josephson junction. *J. Appl. Phys.* **77**, 2582–2590 (1995).
5. Rouse, R., Han, S. & Lukens, J. E. Flux amplification using stochastic superconducting quantum interference devices. *Appl. Phys. Lett.* **66**, 108–110 (1995).
6. Spano, M. L., Wun-Fogle, M. & Ditto, W. L. Experimental observation of stochastic resonance in a magnetoelastic ribbon. *Phys. Rev. A* **46**, R5253–R5256 (1992).
7. Levin, J. E. & Miller, J. P. Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance. *Nature* **380**, 165–168 (1996).
8. Douglass, J. K., Wilkens, L., Pantazelou, E. & Moss, F. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature* **365**, 337–340 (1993).
9. Carr, S. M., Lawrence, W. E. & Wybourne, M. N. Accessibility of quantum effects in mesomechanical systems. *Phys. Rev. B* **64**, 220101 (2001).
10. Nayfeh, A. H. & Mook, D. T. *Nonlinear Oscillations* (Wiley, New York, 1979).
11. Weaver, W., Timoshenko, S. P. & Young, D. H. *Vibration Problems in Engineering* (Wiley, New York, 1990).
12. Badzey, R. L., Zolfagharkhani, G., Gaidarzhay, A. & Mohanty, P. A controllable nanomechanical memory element. *Appl. Phys. Lett.* **85**, 3587–3589 (2004).
13. Benzi, R., Sutera, A. & Vulpiani, A. Stochastic resonance in the Landau–Ginzburg equation. *J. Phys. A* **18**, 2239–2245 (1985).
14. Greywall, D. S., Yurke, B., Busch, P. A., Pargellis, A. N. & Willett, R. L. Evading amplifier noise in nonlinear oscillators. *Phys. Rev. Lett.* **72**, 2992–2995 (1994).
15. Badzey, R. L., Zolfagharkhani, G., Gaidarzhay, A. & Mohanty, P. Temperature dependence of a nanomechanical switch. *Appl. Phys. Lett.* **86**, 023106 (2005).
16. Inchiosa, M. E., Bulsara, A. R. & Gammaitoni, L. Higher-order resonant behaviour in asymmetric nonlinear stochastic systems. *Phys. Rev. E* **55**, 4049–4056 (1997).
17. Inchiosa, M. E. & Bulsara, A. R. dc signal detection via dynamical asymmetry in a nonlinear device. *Phys. Rev. E* **58**, 115–127 (1998).
18. Bulsara, A., Jacobs, E. W., Zhou, T., Moss, F. & Kiss, L. Stochastic resonance in a single neuron model—theory and analog simulation. *J. Theor. Biol.* **152**, 531–555 (1991).
19. Gammaitoni, L., Marchesoni, F., Menichellasaetta, E. & Santucci, S. Multiplicative stochastic resonance. *Phys. Rev. E* **49**, 4878–4881 (1994).
20. McNamara, B. & Wiesenfeld, K. Theory of stochastic resonance. *Phys. Rev. A* **39**, 4854–4869 (1989).
21. Morillo, M. & Gomez-Ordoñez, J. Amplification and distortion of a periodic rectangular driving signal by a noisy bistable system. *Phys. Rev. E* **51**, 999–1003 (1995).
22. Casado-Pascual, J., Gomez-Ordoñez, J. & Morillo, M. Nonlinear stochastic resonance with subthreshold rectangular pulses. *Phys. Rev. E* **69**, 067101 (2004).
23. Gaidarzhay, A., Zolfagharkhani, G., Badzey, R. & Mohanty, P. Evidence for quantized displacement in macroscopic nanomechanical oscillators. *Phys. Rev. Lett.* **94**, 030402 (2005).

Acknowledgements We acknowledge the Nanoscale Exploratory Research (NER) program of the National Science Foundation and the DOD/ARL for the financial support of this research.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.M. (mohanty@physics.bu.edu).

Synthesis and properties of crosslinked recombinant pro-resilin

Christopher M. Elvin¹, Andrew G. Carr¹, Mickey G. Huson², Jane M. Maxwell², Roger D. Pearson¹, Tony Vuocolo¹, Nancy E. Liyou¹, Darren C. C. Wong^{1,3}, David J. Merritt³ & Nicholas E. Dixon⁴

Resilin is a member of a family of elastic proteins that includes elastin, as well as gluten, gliadin, abductin and spider silks. Resilin is found in specialized regions of the cuticle of most insects, providing low stiffness, high strain and efficient energy storage^{1,2}; it is best known for its roles in insect flight^{3,4} and the remarkable jumping ability of fleas^{5,6} and spittle bugs⁷. Previously, the *Drosophila melanogaster* CG15920 gene was tentatively identified as one encoding a resilin-like protein^{8,9} (pro-resilin). Here we report the cloning and expression of the first exon of the *Drosophila* CG15920 gene as a soluble protein in *Escherichia coli*. We show that this recombinant protein can be cast into a rubber-like biomaterial by rapid photochemical crosslinking. This observation validates the role of the putative elastic repeat motif in resilin function. The resilience (recovery after deformation) of crosslinked recombinant resilin was found to exceed that of unfilled synthetic polybutadiene, a high resilience rubber. We believe that our work will greatly facilitate structural investigations into the functional properties of resilin and shed light on more general aspects of the structure of elastomeric proteins. In addition, the ability to rapidly cast samples of this biomaterial may enable its use *in situ* for both industrial and biomedical applications.

The outstanding mechanical properties of resilin were discovered four decades ago during studies of the flight systems of desert locusts and dragonflies^{3,10}. Weis-Fogh proposed a model for elasticity in which randomly coiled, crosslinked polypeptide chains underwent a decrease in entropy upon straining^{10,11}. The chemical crosslinks in resilin from insect joints and tendons occur between tyrosine residues, generating di- and trityrosine^{1,12}. A structural description of the functional properties of the family of elastic proteins has been elusive, but recent studies on elastin^{13–15} suggest that its elasticity is due to the extremely dynamic nature of amorphous hydrophobic domains, which form a kinetically free, random-network polymer.

Naturally crosslinked resilin exhibits two outstanding material properties: it has high rubber efficiency (resilience) and a very high fatigue lifetime. Apart from its role in flight and locomotion, it is also used for other insect functions where efficient energy storage and repetitive movement are required—for example, in the sound-producing organs of cicadas¹⁶ and moths¹⁷, as well as in crustaceans¹⁸. It is also present in some stretchable cuticular structures that do not possess long-range elasticity, including the abdominal wall of physogastric termite queens¹⁹ and in the spermatophore wall of ticks²⁰.

All elastic proteins contain distinct domains, of which at least one is made up of elastic repeat sequences, and they all contain crosslinks between residues in either the non-elastic or elastic domains²¹. To prepare recombinant resilin, we chose to express the first exon of the

Drosophila CG15920 gene, which encodes an amino-terminal domain in the native protein comprising 17 copies of the putative elastic repeat motif, GGRPSDSYGAPGGGN (ref. 8). We cloned and expressed this region (which we refer to as rec1-resilin) as a soluble protein in *E. coli* (Fig. 1a, Methods and Supplementary Information, part 1).

To generate solid, crosslinked rec1-resilin, conditions were established for formation of dityrosine crosslinks between soluble proteins to produce very high molecular weight crosslinked polymeric material. Quantitative crosslinking of soluble rec1-resilin was achieved by use of *Arthromyces ramosus* peroxidase, an enzyme known to catalyse dityrosine formation²². A solid biomaterial formed only from concentrated (>100 mg ml⁻¹) solutions of purified rec1-resilin. This method was of limited value for casting material for physical testing, since the rate of the enzyme-catalysed reaction was uncontrollably fast and the catalase activity of the peroxidase formed oxygen bubbles in the sample. On the other hand, Ru(II)-mediated photo-crosslinking²³ resulted in rapid, quantitative and

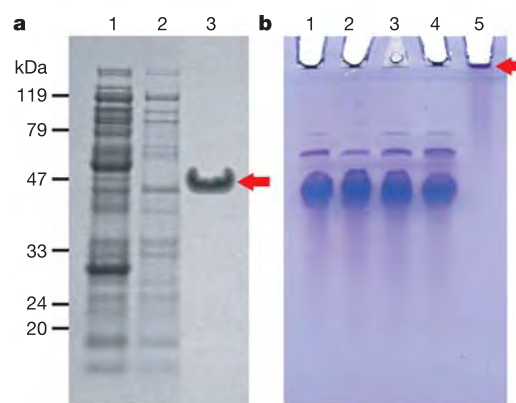


Figure 1 | Purification and crosslinking of soluble recombinant rec1-resilin. **a**, 10% SDS-PAGE of samples from steps in the purification. Lane 1, cleared supernatant from lysed cells; lane 2, flow-through from Q-Sepharose; lane 3, protein eluted from Ni-NTA agarose with a 50–200 mM imidazole gradient (arrow) (see Supplementary Information, part 1.6 for details). **b**, Photo-crosslinking of soluble rec1-resilin analysed on SDS-PAGE. Irradiation (600 W at 15 cm) of a solution of protein (1 mg ml⁻¹) containing [Ru(bpy)₃]²⁺ (200 μM) and ammonium persulphate (APS; 10 mM) produces polymerized rec1-resilin which barely enters a 10% SDS-PAGE gel (arrow). Lane 1, with Ru(II) and APS (no light); lane 2, rec1-resilin only; lane 3, with Ru(II) only; lane 4, with APS only; lane 5, with Ru(II) and APS, 20-s irradiation as described in Supplementary Information, part 1.8.

¹CSIRO Livestock Industries, Queensland Bioscience Precinct, St Lucia 4072, Australia. ²CSIRO Textile and Fibre Technology, PO Box 21, Geelong 3216, Australia. ³School of Integrative Biology, University of Queensland, St Lucia 4072, Australia. ⁴Research School of Chemistry, Australian National University, Canberra 0200, Australia.

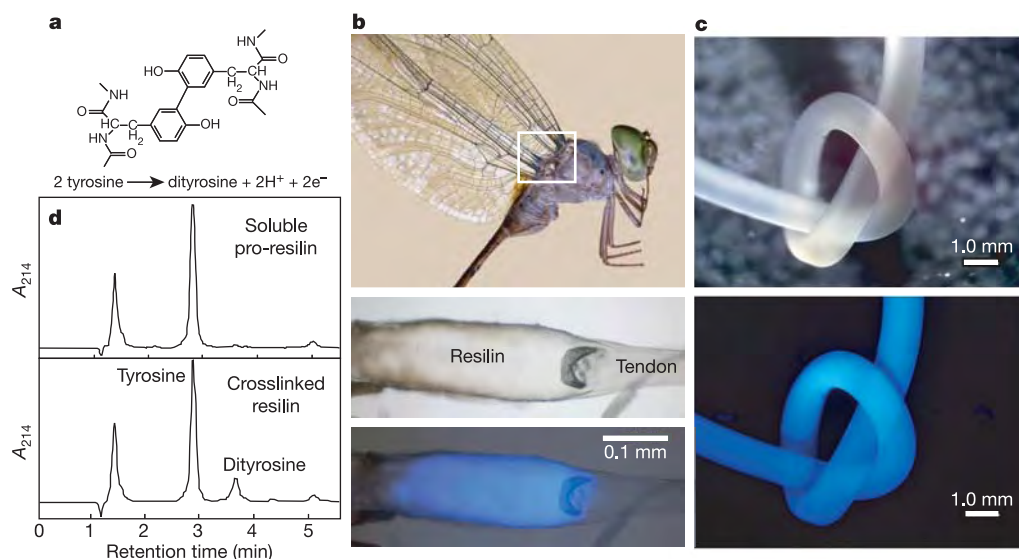


Figure 2 | Dityrosine formation in crosslinked recombinant rec1-resilin.

a, Structure of the dityrosine adduct. **b**, Fluorescence of resilin in the wing tendon from adult dragonfly (*Zygomma* sp.). The lower panels show photomicrographs of the tendon in phosphate-buffered saline under white light and ultraviolet light. **c**, Fluorescence of a moulded rod. A solution of rec1-resilin (200 mg ml⁻¹) containing 2 mM [Ru(bpy)₃]²⁺ and 10 mM persulphate was drawn into a siliconized glass capillary tube (1 mm

diameter) and irradiated with white light. The rod was removed and washed extensively in buffer before photomicrography under white (upper panel) or ultraviolet (lower panel) light. **d**, HPLC analysis of acid-hydrolysed uncrosslinked rec1-resilin and crosslinked rec1-resilin. Samples were analysed by reverse-phase HPLC for dityrosine as described in Supplementary Information, part 1.10. A₂₁₄ is the absorbance at 214 nm.

controllable conversion of the soluble monomer to a high-molecular-weight species (see Fig. 1b and Supplementary Information, parts 1.7 and 1.8). In contrast, recombinant elastin polymers are crosslinked using either γ -irradiation²⁴, lengthy curing using a copper redox system or enzymatic procedures²⁵.

To prepare the solid polymer reproducibly from rec1-resilin solutions, the optimal protein concentration for crosslinking into a hydrogel was determined. The protein was concentrated at 4°C during which a phase separation occurred to produce a denser phase containing ~300 mg ml⁻¹ of rec1-resilin protein. Furthermore, solid hydrogels did not form from solutions at less than 100 mg of protein per ml, using either crosslinking method. We therefore routinely used rec1-resilin at 200 mg ml⁻¹. Various moulded forms of the polymer were cast, physical properties including resilience and stress-strain characteristics were measured, and dityrosine determinations were carried out on this material.

Dityrosine (Fig. 2a) endows natural resilin with pH-dependent blue fluorescence²⁶ on ultraviolet irradiation; the absorption and emission maxima are at 315 and 409 nm, respectively. This is illustrated here in the wing tendon of the dragonfly (Fig. 2b). The resilin contained in this structure is related to rec1-resilin, as evidenced by cross-reaction with an anti-rec1-resilin antibody (D.C.C.W. *et al.*, unpublished observation). The crosslinked rec1-resilin material was similarly fluorescent (Fig. 2c). Amino acid analysis by HPLC (high-performance liquid chromatography) separation after acid hydrolysis showed that the crosslinked rec1-resilin contains dityrosine (Fig. 2d). The extent of dityrosine formation increased with time of irradiation to a maximum near 21% (relative to total tyrosine) after 30 s (Fig. 1a; Supplementary Information, part 2); samples with >17% dityrosine were solid, whereas those with lesser degrees of crosslinking were not. By comparison, Andersen¹² showed that about 25% of tyrosine occurs as the dityrosine dimer in natural locust wing hinge resilin.

Resilience is a measure of a material's ability to recover after deforming under an applied stress. Measurements (Supplementary Information, part 3) on synthetic rubber, native resilin and crosslinked rec1-resilin were made using a scanning probe microscope

(SPM) operated in force mode (Fig. 3a). A tendon from the wing region of the dragonfly was dissected, and a cross-section with exposed resilin was mounted in phosphate-buffered saline such that SPM force measurements could be made. The sample was shown to be 92% resilient. The crosslinked rec1-resilin also displayed negligible hysteresis upon compression (Fig. 3a). A thin rod sample and a strip sample were shown to be 92% and 90% resilient, respectively; this performance is far superior to that of a known low-resilience rubber, chlorobutyl rubber (56%), and even to that of high-resilience polybutadiene rubber (80%) (Fig. 3b). The strip samples were also studied using a tensile tester (Supplementary Information, part 4). The resilience of the biomaterial in phosphate-buffered saline was shown to be 97%, consistent with the SPM studies above. These figures are similar to those reported (90%) for elastin². Furthermore, a strip of crosslinked recombinant resilin could be stretched to over 300% of its original length without breaking (Supplementary Fig. 2).

The resilience of native resilin is very sensitive to changes in both the rate of testing and the extent of deformation, as is the case for elastin². Rec1-resilin, too, like most polymers, exhibits viscoelastic properties. The SPM and tensile tests resulted in the material being deformed at about 5 Hz and 10⁻³ Hz, respectively. In separate experiments on synthetic rubbers, we found that although the absolute value of resilience as measured by SPM might be different from that obtained by more traditional methods, the ranking of the samples remains the same (Supplementary Information, part 3). It should also be noted that rec1-resilin is highly swollen (80% water at equilibrium) and consequently is quite soft. Stress-strain data show it to have a modulus at 100% of about 2.5 kPa, considerably below that of typical unfilled synthetic elastomers (700 kPa)²⁷ and native elastin (1,200 kPa)²⁸ (Supplementary Information, part 4).

From the stress-strain data it is also possible to estimate the crosslink density using the statistical theory of rubber-like elasticity^{28,29} (Supplementary Information, part 4). These data only fit the model during the initial stages of strain (up to 20%), so the analysis was restricted to this portion of the data. If we assume that the density of resilin is 1.33 g cm⁻³, the same as elastin²⁸, and make

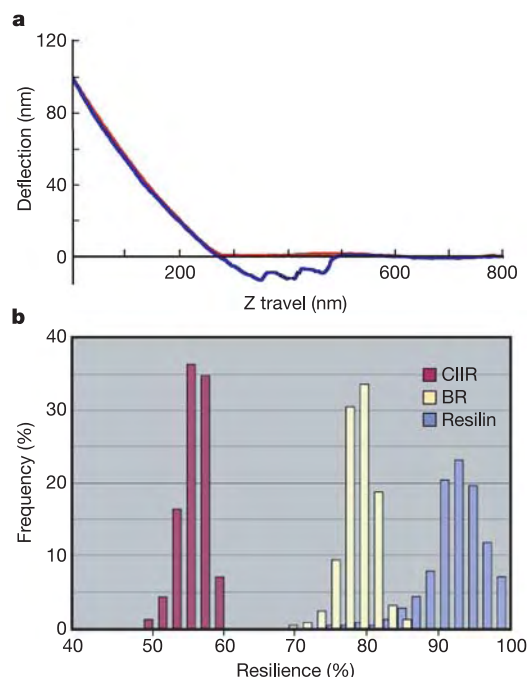


Figure 3 | Elastic properties of crosslinked recombinant resilin. **a**, A single force–extension curve recorded for a sample of crosslinked recombinant resilin. The approach curve is shown in red and the retract curve in blue. **b**, Resilience measurements of elastomers. Samples of chlorobutyl rubber (CIIR), polybutadiene rubber (BR) and crosslinked rec1-resilin were analysed for resilience using the scanning probe microscope as described in Supplementary Information, part 3.

allowance for the fact that the protein elastomer is swollen and also for the effect of chain ends, then the molecular mass between crosslinks (M_c) can be estimated as $14,000 \text{ g mol}^{-1}$. The deduced amino acid sequence of rec1-resilin (molecular weight 28,492) contains 17 tyrosine residues, of which ~20% are converted into dityrosine (Supplementary Fig. 1). From these data, M_c can be estimated as $8,500 \text{ g mol}^{-1}$, in reasonable agreement with the estimate from stress–strain data. This assumes that all dityrosine molecules form effective crosslinks; however, many may be intramolecular, leading to a higher actual M_c .

Time course studies showed that the level of dityrosine crosslinking reached a plateau at 20%, and that the resilience and modulus (inversely proportional to SPM probe penetration) did not increase once the maximum level of dityrosine crosslinking had been achieved (Supplementary Fig. 1).

The physical properties of the resilin polymer are probably due to the three-dimensional amorphous nature of the crosslinked protein matrix and the role of water as plasticizer, as dehydrated crosslinked resilin became very glassy and brittle, and in a partially dry environment (70% relative humidity) it was leathery with poor resilience. It was possible, however, to store the dehydrated material for extended periods and then to recover its normal resilience on rehydration for 2 h in phosphate-buffered saline.

To comment on the lifetime fatigue properties of natural resilin, we investigated the expression of the CG15920 gene in various life stages of *Drosophila*. Total RNA was prepared from cells at each life stage, and first-strand CG15920 complementary DNA was synthesized by reverse transcription (RT) and quantified by real-time quantitative polymerase chain reaction (PCR; Supplementary Information, part 5); data were standardized against expression of the gene for the ribosomal protein RPLPO. The results (Fig. 4) clearly show that CG15920 is expressed only during the pupal stage, when its expression is several thousand-fold greater than in both larval and adult stages. This result is consistent with the developmental profile

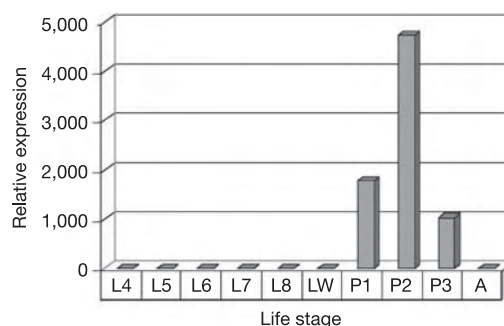


Figure 4 | Real-time quantitative RT-PCR analysis of expression of the *Drosophila* CG15920 gene. Total RNA was prepared from various life stages (L4–L8, larvae day 4 to day 8; P1, newly pupated; P2, contracted light-coloured pupae; P3, dark pupae; A, adult flies) of *Drosophila* (Oregon R) and quantitative real-time RT-PCR carried out on first-strand cDNA. CG15920 messenger RNA levels were normalized to those of the gene for the ribosomal protein RPLPO, and are shown relative to that in the adult (Supplementary Information, part 5).

of insects such as *Drosophila* that undergo complete metamorphosis, laying down their adult cuticle in the pupal stage. Thus, the resilin synthesized at the pupal stage must survive for the lifetime of the adult insect.

If one considers the number of contraction/extension cycles that resilin must accomplish during the course of an insect's life, then, whether it is in the tymbal organ of a cicada producing sound vibrations at 4 kHz (ref. 16) or in a fly flapping its wings at $720,000 \text{ cycles h}^{-1}$ (ref. 30), large numbers are found. In this respect, resilin resembles crosslinked elastin in human arteries, which must also survive for the entire lifetime of the organism². Moreover, when the frog hopper accelerates at about 400g, the resilin-containing tendon releases all of its stored energy in approximately 1 ms (ref. 7). This is similar to the timescale (~3 ms) of contraction/extension that must occur during the wing beat of some dipteran flies³⁰. These very fast contraction rates imply a similarly rapid rate of protein chain movement at the molecular level, consistent with a highly swollen amorphous material.

METHODS

Preparation of crosslinked rec1-resilin. To express the CG15920 gene fragment, genomic DNA from adult *D. melanogaster* was used as template for PCR with primers designed to amplify a 946-base-pair fragment (including codons for an N-terminal His₆ affinity tag) that was inserted into a T7-promoter expression vector. Recombinant rec1-resilin production was induced with IPTG in *E. coli* BL21(DE3)/pLysS and the protein was purified from the cell-free extract by a combination of anion exchange and Ni-NTA affinity chromatography. While the apparent molecular weight appeared to be about 47,000 by SDS–polyacrylamide gel electrophoresis (SDS–PAGE), and sedimentation equilibrium analysis and capillary HPLC in SDS buffer gave values near 23,000, the calculated value is 28,492. To produce hydrogels by photochemical crosslinking²³, solutions of rec1-resilin (200 mg ml^{-1}), 2 mM [Ru(bpy)₃]²⁺ and 10 mM ammonium persulphate in phosphate-buffered saline were irradiated in glass moulds for 10–20 s at 15 cm from the sample with a 600-W white light source. Full details of experimental procedures are given in Supplementary Information.

Received 24 January; accepted 12 July 2005.

- Andersen, S. O. The crosslinks in resilin identified as dityrosine and trityrosine. *Biochim. Biophys. Acta* **93**, 213–215 (1964).
- Gosline, J. et al. Elastic proteins: biological roles and mechanical properties. *Phil. Trans. R. Soc. Lond. B* **357**, 121–132 (2002).
- Weis-Fogh, T. A rubber-like protein in insect cuticle. *J. Exp. Biol.* **37**, 887–907 (1960).
- Gorb, S. N. Serial elastic elements in the damselfly wing: mobile vein joints contain resilin. *Naturwissenschaften* **86**, 552–555 (1999).
- Neville, A. C. & Rothschild, M. Fleas—insects which fly with their legs. *Proc. R. Entomol. Soc. Lond.* **32**, 9–10 (1967).

6. Rothschild, M. & Schlein, J. The jumping mechanism of *Xenopsylla cheopis*. I. Exoskeletal structures and musculature. *Phil. Trans. R. Soc. Lond. B* **271**, 457–490 (1975).
7. Burrows, M. Biomechanics: frog hopper insects leap to new heights. *Nature* **424**, 509 (2003).
8. Ardell, D. H. & Andersen, S. O. Tentative identification of a resilin gene in *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* **31**, 965–970 (2001).
9. Andersen, S. O. in *Elastomeric Proteins* (eds Shewry, P. R., Tatham, A. S. & Bailey, A. J.) 259–278 (Cambridge Univ. Press, Cambridge, UK, 2002).
10. Weis-Fogh, T. Molecular interpretation of the elasticity of resilin, a rubber-like protein. *J. Mol. Biol.* **3**, 520–531 (1961).
11. Weis-Fogh, T. in *The Cell and the Organism* (eds Ramsay, J. A. & Wigglesworth, V. B.) 283–300 (Cambridge Univ. Press, Cambridge, UK, 1961).
12. Anderson, S. O. Covalent cross-links in a structural protein, resilin. *Acta Physiol. Scand. Suppl.* **263**, 1–81 (1966).
13. Li, B. & Daggett, V. Molecular basis for the extensibility of elastin. *J. Muscle Res. Cell Motil.* **23**, 561–573 (2002).
14. Rousseau, R., Schreiner, E., Kohlmeyer, A. & Marx, D. Temperature-dependent conformational transitions and hydrogen-bond dynamics of the elastin-like octapeptide GVG(VPGVG): A molecular-dynamics study. *Biophys. J.* **86**, 1393–1407 (2004).
15. Pometun, M. S., Chekmenev, E. Y. & Wittebort, R. J. Quantitative observation of backbone disorder in native elastin. *J. Biol. Chem.* **279**, 7982–7987 (2004).
16. Young, D. & Bennet-Clark, H. C. The role of the tymbal in cicada sound production. *J. Exp. Biol.* **198**, 1001–1019 (1995).
17. Skals, N. & Surlykke, A. Sound production by abdominal tymbal organs in two moth species: The green silver-line and the scarce silver-line (Noctuoidea: Nolidae: Chloephorinae). *J. Exp. Biol.* **202**, 2937–2949 (1999).
18. Andersen, S. O. & Weis-Fogh, T. Resilin: a rubber-like protein in arthropod cuticle. *Adv. Insect Physiol.* **2**, 1–65 (1964).
19. Varman, A. R. Resilin in the cuticle of physogastric queen termites. *Experientia* **36**, 564 (1980).
20. Singaravelu, G. Occurrence of elastic protein resilin in the spermatophore walls of a tick *Haemaphysalis intermedia*. *Natl Acad. Sci. Lett. (India)* **14**, 147–149 (1991).
21. Tatham, A. S. & Shewry, P. R. Comparative structures and properties of elastic proteins. *Phil. Trans. R. Soc. Lond. B* **357**, 229–234 (2002).
22. Malencik, D. A. & Anderson, S. R. Dityrosine formation in calmodulin: cross-linking and polymerization catalyzed by *Arthromyces* peroxidase. *Biochemistry* **35**, 4375–4386 (1996).
23. Fancy, D. A. & Kodadek, T. Chemistry for the analysis of protein–protein interactions: rapid and efficient cross-linking triggered by long wavelength light. *Proc. Natl Acad. Sci. USA* **96**, 6020–6024 (1999).
24. Urry, D. W. et al. in *Elastomeric Proteins* (eds Shewry, P. R., Tatham, A. S. & Bailey, A. J.) 54–93 (Cambridge Univ. Press, Cambridge, UK, 2002).
25. Keeley, F. W., Bellingham, C. M. & Woodhouse, K. A. Elastin as a self-organising biomaterial: use of recombinantly expressed human elastin polypeptides as a model system for investigations of structure and self-assembly of elastin. *Phil. Trans. R. Soc. Lond. B* **357**, 185–189 (2002).
26. Neff, D., Frazier, S. F., Quimby, L., Wang, R. T. & Zill, S. Identification of resilin in the leg of cockroach, *Periplaneta americana*: confirmation by a simple method using pH dependence of UV fluorescence. *Arthropod Struct. Dev.* **29**, 75–83 (2000).
27. Cowie, J. M. G. *Polymers: Chemistry and Physics of Modern Materials* (Int. Textbook Co. Ltd, Aylesbury, 1973).
28. Aaron, B. B. & Gosline, J. M. Elastin as a random-network elastomer: a mechanical and optical analysis of single elastin fibers. *Biopolymers* **20**, 1247–1260 (1980).
29. Treloar, L. R. G. *The Physics of Rubber Elasticity* (Clarendon, Oxford, 1975).
30. Lehmann, F. O. & Dickinson, M. H. The production of elevated flight force compromises manoeuvrability in the fruit fly *Drosophila melanogaster*. *J. Exp. Biol.* **204**, 627–635 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Abbenante for preparative scale reverse-phase HPLC purification of dityrosine and for mass spectrometry. The technical assistance of L. Conlan is acknowledged for HPLC analysis of dityrosine. We are grateful to A. Brownlee and R. Tellam for critical reading of the manuscript. This work was supported by a CSIRO Nanotechnology Emerging Sciences Initiative grant.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.M.E. (chris.elvin@csiro.au).

Climatic controls on central African hydrology during the past 20,000 years

Enno Schefuß^{1†}, Stefan Schouten² & Ralph R. Schneider³

Past hydrological changes in Africa have been linked to various climatic processes, depending on region and timescale. Long-term precipitation changes in the regions of northern and southern Africa influenced by the monsoons are thought to have been governed by precessional variations in summer insolation^{1,2}. Conversely, short-term precipitation changes in the northern African tropics have been linked to North Atlantic sea surface temperature anomalies, affecting the northward extension of the Intertropical Convergence Zone and its associated rainbelt^{3,4}. Our knowledge of large-scale hydrological changes in equatorial Africa and their forcing factors is, however, limited⁵. Here we analyse the isotopic composition of terrigenous plant lipids, extracted from a marine sediment core close to the Congo River mouth, in order to reconstruct past central African rainfall variations and compare this record to sea surface temperature changes in the South Atlantic Ocean. We find that central African precipitation during the past 20,000 years was mainly controlled by the difference in sea surface temperatures between the tropics and subtropics of the South Atlantic Ocean, whereas we find no evidence that changes in the position of the Intertropical Convergence Zone had a significant influence on the overall moisture availability in central Africa. We conclude that changes in ocean circulation, and hence sea surface temperature patterns, were important in modulating atmospheric moisture transport onto the central African continent.

Life and vegetation in low-latitude continental regions depend primarily on the availability of water. It is, therefore, of crucial significance to improve our understanding of tropical hydrological changes and their causes. Continental records of past precipitation changes in central Africa indicate a dry Last Glacial Maximum^{5,6} (LGM), 23,000–18,000 years before present (yr BP), followed by an early deglacial humidity increase at 17,000 yr BP⁵. During the last deglaciation and the Holocene, several dry spells (thought to have been caused by North Atlantic cold sea surface temperature (SST) anomalies⁴) occurred in the northern tropics⁵, increasing the inter-hemispheric SST gradient and displacing the Intertropical Convergence Zone (ITCZ) and its associated rainfall belt southward³. In central Africa, precipitation maxima occur twice a year, when the ITCZ moves across⁷. Present-day meteorological observations suggest a relation between central African rainfall and the meridional South Atlantic SST pattern, indicating a negative correlation of rainfall with equatorial Atlantic SSTs⁸ and a positive correlation with subtropical South Atlantic SSTs^{8,9}.

To better estimate past large-scale hydrological changes in central Africa and to study their link to Atlantic SST distributions, we used alkenone-derived SST reconstructions¹⁰ in combination with hydrogen isotopic analyses of sedimentary long-chain *n*-alkanes¹¹. The latter compounds derive from the epicuticular wax layer of terrestrial

higher plants¹² and thus are direct recorders of continental climate conditions. Terrigenous and marine compounds were extracted from marine sediment core GeoB 6518-1 (05° 35.3' S, 11° 13.3' E, 962 m water depth, Fig. 1), recovered close to the Congo River. The catchment of the Congo River (about 3.8×10^6 km², Fig. 1) integrates equatorial African rainfall changes, thus enabling the simultaneous reconstruction of continental and oceanic climate changes for the Holocene and the last deglaciation with pronounced climatic changes, particularly in high latitudes¹³ (Fig. 2a).

Hydrogen isotope compositions (δD in ‰ VSMOW) of plant-wax lipids are similar for plants using different photosynthetic pathways, that is, C₃ and C₄ plants¹⁴. Large isotopic changes are thus primarily controlled by environmental conditions, in the sense that moisture loss by evapo-transpiration leads to deuterium enrichment in plants¹⁵. Additionally, soil waters become deuterium-enriched by evaporation under more arid conditions. δD values of terrestrial plant waxes therefore reflect the precipitation–evaporation balance. Moreover, the isotopic composition of precipitation in

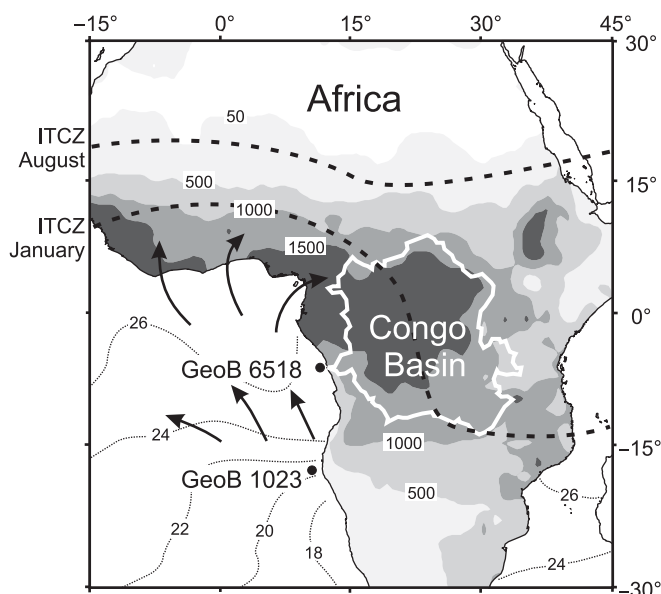


Figure 1 | Overview map showing the main climatological features of Africa and the adjacent Atlantic Ocean. Shown are core locations (filled circles), the extent of the Congo River catchment (white outline), the mean annual precipitation rate (in mm yr⁻¹, shading, based on ref. 30), the mean January and August positions of the ITCZ (dashed lines), the schematic mean annual wind pattern over the Atlantic Ocean (curved arrows), and the annual mean SST field²¹ (contourlines).

¹DFG-Research Center Ocean Margins (RCOM) at the University of Bremen, Klagenfurter Strasse, 28359 Bremen, Germany. ²Royal Netherlands Institute for Sea Research (NIOZ), Landsdiep 4, 1790 AB Den Burg, Texel, The Netherlands. ³Institut für Geowissenschaften, Christian-Albrechts-Universität zu Kiel, Ludwig-Meyn-Strasse 10, 24118 Kiel, Germany. [†]Present address: Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

low-latitude regions is coupled to humidity, as higher tropical precipitation rates lead to decreased δD values of rain, and vice versa¹⁶. Temperature effects on the isotopic composition of precipitation are restricted to higher latitudes¹⁷. Overall, δD variations of tropical plant waxes predominantly reflect continental rainfall fluctuations.

δD values of the n -C₂₉ alkane, the dominant homologue of the sedimentary plant waxes in core GeoB 6518-1, range from -123‰

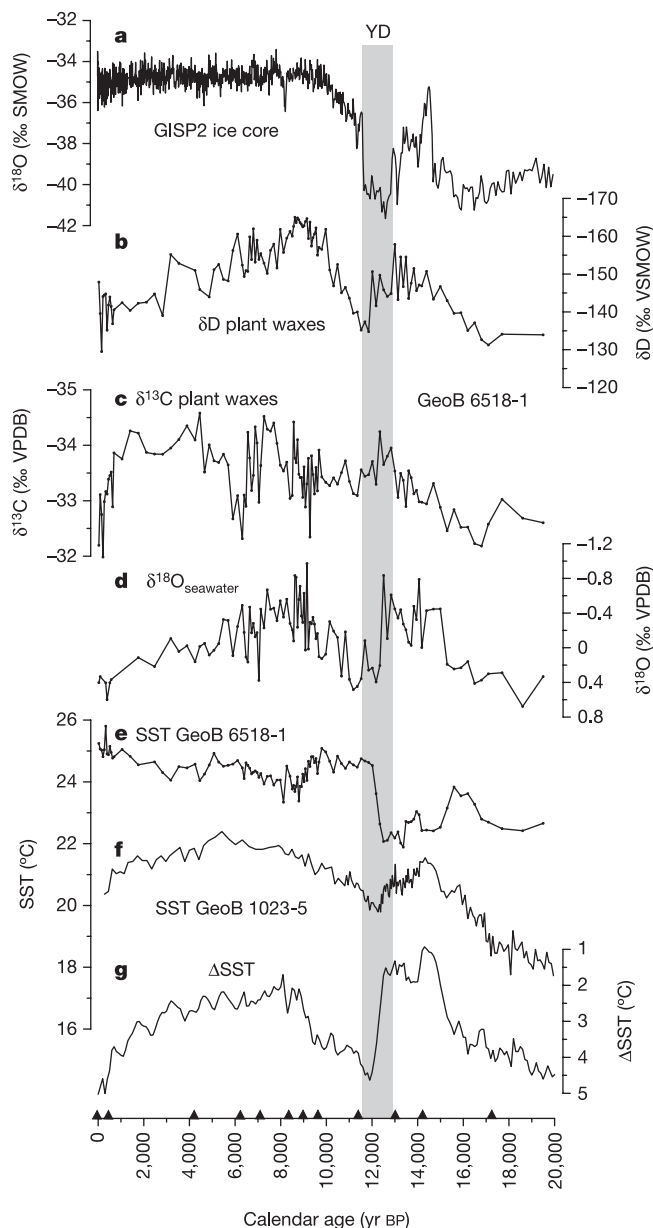


Figure 2 | Deglacial-Holocene records of northern high-latitude and central African climate and SST variations in the tropical and subtropical Atlantic. Shown is the GISP2 $\delta^{18}\text{O}$ record reflecting Greenland air temperatures¹³ (a), the plant-wax δD record from GeoB 6518-1 reflecting large-scale central African precipitation changes (b), the plant-wax $\delta^{13}\text{C}$ record from GeoB 6518-1 reflecting relative contributions from C₃ and C₄ plants (c), the $\delta^{18}\text{O}$ record of Congo River discharge (d), the SST development in the equatorial Atlantic (GeoB 6518-1) (e) and the subtropical South Atlantic (GeoB 1023-5, ref. 26) (f), and the tropical–subtropical sea surface temperature gradient (ΔSST) in the South Atlantic (g). ^{14}C -AMS dates of GeoB 6518-1 (Table 1) are indicated by triangles on the lower axis. The grey bar indicates the Younger Dryas period (YD; 12,900–11,600 yr BP).

to -163‰ (VSMOW). We corrected these values for global isotopic changes of ocean waters due to the decay of continental ice sheets¹⁸, resulting in a slightly smaller range from -131‰ to -165‰ (Fig. 2b). Highest values, indicating least precipitation, occurred around 17,100 yr BP. Until about 13,000 yr BP, δD values continuously decreased, that is, rainfall increased. At 12,900 yr BP, the beginning of the Younger Dryas period (YD), δD values abruptly increased by more than 20‰, suggesting a rapid aridification, coinciding with lowered lake levels throughout equatorial Africa⁵. Following the YD, δD values gradually decreased again, until minimum values were reached around 9,000 yr BP, suggesting highest precipitation rates and supporting continental studies inferring very humid conditions for the early to mid-Holocene^{5,6}. Afterwards, gradually increasing δD values imply aridification during the late Holocene, in accordance with continental palaeoenvironmental reconstructions^{5,6}. This agreement between the hydrological changes reconstructed here and those of continental records rules out a major influence of continental temperatures on the δD values, as elevated Holocene temperatures would have led to deuterium-enriched precipitation¹⁷, rather than the observed depletion in δD .

In contrast to hydrogen isotopes, stable carbon isotopic compositions of plant waxes reflect different vegetation types. Leaf waxes from C₃ plants (for example, rainforest trees¹⁹) have stable carbon isotope compositions ($\delta^{13}\text{C}$ in ‰ VPDB) as low as -36‰ (ref. 14), whereas those from C₄ plants (for example, tropical grasses¹⁹) are as high as -20‰ (ref. 14). Environmental factors, such as humidity and temperature, influence $\delta^{13}\text{C}$ values of plants to a lesser degree. The $\delta^{13}\text{C}$ value of the n -C₂₉ alkane varies in a narrow range of -32.0‰ and -34.6‰ (Fig. 2c), indicating that the Congo River mainly drains rainforest-vegetated areas. The small range of the $\delta^{13}\text{C}$ variations suggests only minor changes in the relative proportions of C₃ and C₄ plant-derived wax lipids, confirming that the relatively large changes in δD were independent of changes in vegetation type. The absence of δD variations during times of rapid changes in the plant-wax $\delta^{13}\text{C}$ values (for example, around 6,000 yr BP) supports this conclusion. Palaeovegetation data indicate that C₄ plants were more widespread in equatorial Africa during the LGM, were partly replaced by C₃ plants during deglaciation, and flourished again during the YD⁶. C₃ plants dominated the equatorial vegetation throughout the early to mid-Holocene⁶. The short-term $\delta^{13}\text{C}$ increases during the Holocene (Fig. 2c), however, cannot indicate the expansion of C₄ grasslands in central Africa, as they coincide with high rainfall estimates (Fig. 2b). We therefore suggest that they might reflect source area changes within the Congo catchment, possibly connected to ITCZ movements. Inside the catchment, areas with elevated relative C₄ plant abundance exist¹⁹, which could have supplied more C₄ plant-derived lipids to the marine sediments when receiving more rainfall. As plants take up the isotopic rainwater signal rather rapidly, they would adjust their hydrogen isotopic composition, whereas they would retain their carbon isotopic signature. A spread of forest into grassland regions would require much longer times than the uptake of a changed hydrogen isotopic signal of precipitation by plants.

The use of hydrogen isotopic compositions of fluvially-transported lipids for continental rainfall reconstructions is further validated by changes recorded in the oxygen isotope compositions ($\delta^{18}\text{O}$ in ‰ VPDB) of planktonic foraminifera from the same samples. The $\delta^{18}\text{O}_{\text{carbonate}}$ values have been corrected for the effects of post-glacial ice-sheet decay¹⁸ on $\delta^{18}\text{O}_{\text{seawater}}$ and calcification temperatures²⁰, using alkenone-derived SSTs (Fig. 2e). The residual $\delta^{18}\text{O}_{\text{seawater}}$ values (Fig. 2d) thus reflect isotopic changes in the surface waters, determined by the variable ^{18}O -depleted freshwater discharge of the Congo River. The $\delta^{18}\text{O}_{\text{seawater}}$ changes closely resemble the δD record of the plant waxes in relative amplitude and timing, suggesting a close connection between continental precipitation and runoff.

SST changes in the moisture source region may also change the

deuterium content of vapour by non-equilibrium processes during evaporation, leading to deuterium excess in precipitation¹⁷. Therefore, we compare the δD record to the alkenone-derived SST record from GeoB 6518-1 (Fig. 2e). Modern SSTs in the Congo River plume closely follow the seasonal SST variation in the eastern equatorial Atlantic²¹. Core-top alkenone-based SST from the Congo fan do not deviate from the global calibration¹⁰. Glacial–interglacial SST records from the Congo fan covary strongly with SST records from outside the Congo plume, exhibiting identical temperature ranges²². The GeoB 6518-1 SST record also largely resembles a deglacial–Holocene intermediate-depth temperature record from the Angolan coast²³. Therefore, we consider the GeoB 6518-1 SST record to reflect the larger-scale SST development in the eastern equatorial Atlantic. SSTs slowly rose by 1.6 °C from the LGM until about 15,500 yr BP and then decreased to values equivalent to those during the LGM. During the YD, SSTs rapidly rose by about 2.5 °C, comparable to the 1–3 °C intermediate-depth warming²³, presumably caused by a reduced thermohaline circulation^{23,24}. Afterwards, SSTs remained generally high until the present. Since the effect of SST changes on the isotopic composition of the initial moisture by deuterium excess is only ~1‰ per °C (ref. 25), its effect was less than 4‰. Therefore, δD variations of the plant waxes were mainly controlled by precipitation changes.

The reconstructed past rainfall variations in central Africa rule out a forcing by precessional insolation changes only. Owing to their opposite effect on monsoonal precipitation in both hemispheres, any insolation-driven precipitation increase north of the Equator was accompanied by a decrease south of it¹. Additionally, increased Northern Hemisphere summer insolation during the early to mid-Holocene would have actually driven the northernmost ITCZ position away from the catchment¹ and thus decreased precipitation in central Africa, contrary to the reconstructed high precipitation rates during that time. The aridification during the YD also contradicts a control of central African rainfall by the interhemispheric SST gradient. Low SST in the subtropical North Atlantic and high SST in the equatorial Atlantic during the YD would have led to a generally more southward position of the ITCZ³, which would have actually increased rainfall in the Congo catchment. Precessional insolation changes and North Atlantic SST anomalies thus cannot explain the observed large-scale central African precipitation changes. The δD record (Fig. 2b) also does not covary with the equatorial SST record from GeoB 6518-1 (Fig. 2e). From 20,000 to 13,000 yr BP, when SSTs varied between 22 and 24 °C, δD values decreased by about 20‰,

whereas from 12,000 yr BP onwards, when SSTs remained relatively stable, δD values decreased by about 30‰ and rose afterwards by the same amplitude. This implies that central African precipitation was also not solely controlled by equatorial SSTs⁸ via the land–ocean pressure gradient. The only exception is the sudden SST rise around 12,500 yr BP, which corresponds to a precipitation decrease.

Taking into account the present-day association of central African rainfall with basin-wide South Atlantic SST distributions^{8,9}, we relate the continental rainfall changes to the meridional South Atlantic SST gradient. Therefore, the tropical SST record is compared to a subtropical South Atlantic SST record²⁶ (GeoB 1023-5, 17° 09.5' S, 11° 00.5' E, 1,978 m water depth; Figs 1, 2f). This SST record again rules out a forcing of central African rainfall by ITCZ movements. The low subtropical SSTs during the YD²⁶ that occurred also in the Southern Hemisphere would have resulted in an equatorward displacement of the southernmost ITCZ position, which cannot explain the lower precipitation in central Africa at this time. The tropical–subtropical SST gradient (Fig. 2g), however, corresponds to a sea-level pressure gradient, which in turn controls the intensity of the Southern Hemisphere trade winds²⁷, counteracting the inflow of moist air from the Atlantic Ocean onto the African continent. Central African precipitation was high when this gradient was low, that is, when the equatorial Atlantic was relatively cool and the subtropical South Atlantic was relatively warm, leading to weak trade winds, such as from 15,000 to 12,500 yr BP and during the early to mid-Holocene (Fig. 2b, g). The reversed situation—relatively cold subtropics and relatively warm tropics—corresponded to stronger trade winds and arid periods in equatorial Africa. The tropical–subtropical SST gradient in the South Atlantic thus controlled large-scale central African hydrological changes over the past 20,000 yr.

It must be noted that our records do not rule out past ITCZ movements or their effect on precipitation changes on a regional scale. However, this process is unable to explain the observed large-scale precipitation changes in central Africa. Unlike rainfall variations in northern and southern Africa, past precipitation changes in central Africa were primarily controlled by the tropical–subtropical SST gradient in the South Atlantic. Millennial-scale changes in ocean circulation^{23,24} apparently had a major effect on central African hydrology by modulating atmospheric moisture transport onto the continent.

METHODS

Radiocarbon chronology. Fifteen ¹⁴C-AMS dates on mixed planktonic foraminifera fractions containing *Globigerinoides ruber* (white), *Globigerinoides sacculifer* and *Orbulina universa* provide the core chronology (Table 1).

Lipid extraction and analyses. About 5 g dried and ground sediment samples were mixed with known amounts of standards and ultrasonically extracted (3 ×) using successively less polar solvent mixtures (CH₃OH, CH₃OH/CH₂Cl₂ 1:1, CH₂Cl₂). Extracts of each sample were combined and rotary-evaporated to near dryness. Extracts were taken up in CH₂Cl₂, desalted with double-distilled H₂O, dried with anhydrous Na₂SO₄ and extracted again with CH₂Cl₂ (3 ×). Elemental sulphur was removed by stirring overnight with activated copper in CH₂Cl₂. Extracts were then cleaned by elution with CH₂Cl₂ over silica columns (Varian Bond elute cartridges). To remove wax esters, extracts were hydrolysed with 1 N KOH in methanol (CH₃OH/H₂O 9:1) at 80 °C for 2 h. Lipids were extracted by partitioning into hexane and dried under N₂. Hydrocarbon fractions were separated by column-chromatography over activated Al₂O₃ eluted with hexane. More polar lipids were eluted from the column with CH₃OH/CH₂Cl₂ (1:1). Saturated hydrocarbon fractions were obtained by column chromatography over AgNO₃-impregnated silica eluted with hexane. Lipid fractions were analysed on a Hewlett Packard 5890 series II gas chromatograph using flame ionisation detection (FID). Quantification of compounds was performed by peak area integration in FID chromatograms relative to standards. The relative abundances of the C_{37:3} and C_{37:2} alkenones were converted into SST estimates using the calibration of ref. 10. Analytical precision, based on duplicate sediment extractions, was always better than 0.3 °C.

Lipid hydrogen and carbon isotopes. Compound-specific hydrogen isotope compositions of *n*-alkanes were determined by gas chromatography–isotope ratio monitoring–mass spectrometry (GC-irm-MS) with a Thermo Electron

Table 1 | ¹⁴C-AMS dates used for the chronology of GeoB 6518-1

Sample	Core depth (cm)	¹⁴ C age (¹⁴ C yr BP)	Error (¹⁴ C yr)	Calendar age range (1σ) (yr BP)	Calendar age (yr BP)
KIA23160	10	160	±90	*	*
KIA23161	58	1,150	+55/–50	649–736	693
KIA23162	108	4,175	±45	4,175–4,319	4,247
KIA23163	158	5,910	±45	6,277–6,357	6,317
KIA23164	208	6,605	±60	7,049–7,201	7,125
KIA23165	255.5	8,040	±55	8,412–8,543	8,478
KIA23166	305.5	8,670	+80/–70	8,985–9,115	9,050
KIA23167	355.5	9,130	+90/–80	9,562–9,857	9,710
KIA23252	405.5	10,515	±60	11,324–11,560	11,442
KIA23168	455.5	11,630	±80	12,971–13,189	13,080
KIA23251	505.5	12,780	±70	14,101–14,375	14,238
KIA23249	555.5	14,900	±80	16,989–17,511	17,250
KIA23248	605.5	22,690	±170	†	26,243†
KIA23247	655.5	29,000	+360/–340	†	33,413†
KIA23245	705.5	35,840	+830/–750	†	40,915†

Conversion of radiocarbon ages to calendar ages was done using the program CALIB 4.4²⁸ using the marine calibration and applying a reservoir age of 400 yr.

*The uppermost sample could not be calibrated to a calendar age owing to the influence of bomb-derived radiocarbon. Therefore, we set the core top to zero age.

†The radiocarbon ages of the three lowermost samples exceed the calibration range. Therefore, these dates have been converted to calendar ages using Bard's glacial polynomial²⁹.

Delta plus XL mass spectrometer using high temperature conversion. H₂ gas with known isotopic composition was used as a reference and an internal standard (squalane) was monitored during analyses. Analyses were done at least in duplicate. The average δD value of the internal standard during analyses was $-169.7 \pm 4.2\text{‰}$ VSMOW, close to the offline-determined value of $-170.0 \pm 4.0\text{‰}$. Compound-specific stable carbon isotope composition of *n*-alkanes was measured by GC-irm-MS with a Finnigan MAT 252 mass spectrometer. CO₂ gas with known isotopic composition was used as reference. Analyses were done at least in duplicate. Reproducibility of $\delta^{13}C$ values was always better than 0.5‰ (VPDB).

Foraminifera oxygen isotopes. For oxygen isotope analyses, about 30 specimens of *G. ruber* (white) ($>150\mu m$) were measured using a Finnigan MAT 251 mass spectrometer equipped with an automatic carbonate preparation device. Internal precision, based on replicates of a laboratory standard, was better than 0.07‰ (VPDB).

Received 13 December 2004; accepted 16 June 2005.

- Kutzbach, J. E. & Liu, Z. Response of the African monsoon to orbital forcing and ocean feedbacks in the middle Holocene. *Science* **278**, 440–444 (1997).
- Partridge, T. C., DeMenocal, P. B., Lorentz, S. A., Paiker, M. J. & Vogel, J. C. Orbital forcing of climate over South Africa: a 200,000 year rainfall record from the Pretoria Saltpan. *Quat. Sci. Rev.* **16**, 1125–1133 (1997).
- Hastenrath, S. Decadal-scale changes of the circulation in the tropical Atlantic sector associated with Sahel drought. *Int. J. Climatol.* **10**, 459–472 (1990).
- Street-Perrott, F. A. & Perrott, R. A. Abrupt climate fluctuations in the tropics: the influence of Atlantic Ocean circulation. *Nature* **343**, 607–612 (1990).
- Gasse, F. Hydrological changes in the African tropics since the Last Glacial Maximum. *Quat. Sci. Rev.* **19**, 189–211 (2000).
- Maley, J. & Brenac, P. Vegetation dynamics, palaeoenvironments and climatic changes in the forests of western Cameroon during the last 28,000 years B.P. *Rev. Paleobot. Palynol.* **99**, 157–187 (1998).
- Sanga-Ngoie, K. & Fukuyama, K. Interannual and long-term climate variability over the Zaire River Basin during the last 30 years. *J. Geophys. Res.* **101**, 21351–21360 (1996).
- Camberlin, P., Janicot, S. & Pocard, I. Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO. *Int. J. Climatol.* **21**, 973–1005 (2001).
- Nicholson, S. E. & Entekhabi, D. Rainfall variability in equatorial and southern Africa: relationships with sea surface temperatures along the southwestern coast of Africa. *J. Clim. Appl. Meteorol.* **26**, 561–578 (1987).
- Müller, P. J., Kirst, G., Ruhland, G., von Storch, I. & Rosell-Mele, A. Calibration of the alkenone paleotemperature index U_K^{27} based on core-tops from the eastern South Atlantic and the global ocean (60°N–60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772 (1998).
- Sessions, A. L., Burgoyne, T. W., Schimmelmann, A. & Hayes, J. M. Fractionation of hydrogen isotopes in lipid biosynthesis. *Org. Geochem.* **30**, 1193–1200 (1999).
- Eglinton, G. & Hamilton, R. J. Leaf epicuticular waxes. *Science* **156**, 1322–1335 (1967).
- Stuiver, M. & Grootes, P. M. GISP2 oxygen isotope ratios. *Quat. Res.* **53**, 277–284 (2000).
- Chikaraishi, Y., Naraoka, H. & Poulson, S. R. Hydrogen and carbon isotopic fractionations of lipid biosynthesis among terrestrial (C3, C4 and CAM) and aquatic plants. *Phytochemistry* **65**, 1369–1381 (2004).
- Sauer, P. E., Eglinton, T. I., Hayes, J. M., Schimmelmann, A. & Sessions, A. L. Compound-specific D/H ratios of lipid biomarkers from sediments as a proxy for environmental and climatic conditions. *Geochim. Cosmochim. Acta* **65**, 213–222 (2001).
- Jouzel, J., Hoffmann, G., Koster, R. D. & Masson, V. Water isotopes in precipitation: data/model comparison for present-day and past climates. *Quat. Sci. Rev.* **19**, 363–379 (2000).
- Dansgaard, W. Stable isotopes in precipitation. *Tellus* **16**, 436–468 (1964).
- Fairbanks, R. G., Charles, C. D. & Wright, J. D. in *Radiocarbon After Four Decades: An Interdisciplinary Perspective* (eds Taylor, R. E., Long, A. & Kra, R.) 473–500 (Springer, New York, 1992).
- Still, C. J., Berry, J. A., Collatz, G. J. & DeFries, R. S. Global distribution of C₃ and C₄ vegetation: Carbon cycle implications. *Glob. Biogeochem. Cycles* **17**, doi:10.1029/2001GB001807 (2003).
- Bemis, B., Spero, H., Bijma, J. & Lea, D. Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. *Paleoceanography* **13**, 150–160 (1998).
- Levitus, S. & Boyer, T. P. *World Ocean Atlas 1994 Vol. 4, Temperature* (NOAA Atlas NESDIS 4, US Dept of Commerce, Washington DC, 1994).
- Schneider, R. R., Müller, P. J. & Ruhland, G. Late Quaternary surface circulation in the east equatorial South Atlantic: Evidence from alkenone sea surface temperatures. *Paleoceanography* **10**, 197–219 (1995).
- Rühlemann, C. *et al.* Intermediate depth warming in the tropical Atlantic related to weakened thermohaline circulation: Combining paleoclimate data and modeling results for the last deglaciation. *Paleoceanography* **19**, doi:10.1029/2003PA000948 (2004).
- Manabe, S. & Stouffer, R. J. Coupled ocean-atmosphere model response to freshwater input: Comparison to Younger Dryas event. *Paleoceanography* **12**, 321–336 (1997).
- Steig, E. C. No two latitudes alike. *Science* **293**, 2015–2016 (2001).
- Kim, J.-H., Schneider, R. R., Mulitza, S. & Müller, P. J. Reconstruction of SE trade-wind intensity based on sea-surface temperature gradients in the Southeast Atlantic over the last 25 kyr. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017557 (2003).
- Lindzen, R. S. & Nigam, S. On the role of sea surface temperature gradients in forcing low level winds and convergence in the tropics. *J. Atmos. Sci.* **44**, 2418–2436 (1987).
- Stuiver, M., Reimer, P. J. & Braziunas, T. F. High-precision radiocarbon age calibration for terrestrial and marine samples. *Radiocarbon* **40**, 1127–1151 (1998).
- Bard, E. Geochemical and geophysical implications of the radiocarbon calibration. *Geochim. Cosmochim. Acta* **62**, 2025–2038 (1998).
- Nicholson, S. E. The nature of rainfall variability over Africa on time scales of decades to millennia. *Glob. Planet. Change* **26**, 137–158 (2000).

Acknowledgements We thank L. M. Dupont, T. I. Eglinton, G. Lohmann and S. Mulitza for discussions. We are grateful to W. Bevern, A. Jaeschke, M. Kienhuis, M. Segl and S. P. Sylva for analytical assistance. P. M. Grootes and staff of the Leibniz-Labor for Radiometric Dating and Isotope Research, Kiel provided ¹⁴C-AMS datings. This research was supported by the Deutsche Forschungsgemeinschaft.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.S. (eschefuss@whoi.edu).

The earliest dromaeosaurid theropod from South America

Peter J. Makovicky¹, Sebastián Apesteguía^{2,4} & Federico L. Agnolín³

The evolutionary history of Maniraptora, the clade of carnivorous dinosaurs that includes birds and the sickle-clawed Dromaeosauridae, has hitherto been largely restricted to Late Jurassic and Cretaceous deposits on northern continents. The stunning Early Cretaceous diversity of maniraptorans from Liaoning, China^{1–3}, coupled with a longevity implied by derived Late Jurassic forms such as *Archaeopteryx*, pushes the origins of maniraptoran lineages back to Pangaeian times and engenders the possibility that such lineages existed in Gondwana. A few intriguing, but incomplete, maniraptoran specimens have been reported from South America^{4–8}, Africa⁹ and Madagascar¹⁰. Their affinities remain contested^{11–13}, however, and they have been interpreted as biogeographic anomalies relative to other faunal components of these land-masses. Here we describe a near-complete, small dromaeosaurid that is both the most complete and the earliest member of the Maniraptora from South America, and which provides new evidence for a unique Gondwanan lineage of Dromaeosauridae with an origin predating the separation between northern and southern landmasses.

The new dromaeosaurid was discovered at 'La Buitrera', an extensive fossiliferous locality in the Upper Cretaceous Candeleros Fm. of Río Negro Province (Neuquén Basin) (Fig. 1), which yields abundant and exquisitely preserved three-dimensional skeletons of small tetrapods¹⁴. The 'La Buitrera' fauna is biased towards articulated microvertebrate and mesovertebrate specimens, including abundant spheonodontids¹⁵, araripesuchid crocodyliforms¹⁶, primitive limbed snakes and mammals, including dryolestoids. Dinosaurs are rare, but nearby Candeleros exposures have produced abundant remains of large dinosaurs including carcharodontosaurids¹⁷, rebbachisaurids¹⁸ and basal titanosaurs. In addition to being the earliest undoubted maniraptoran collected from Gondwanan continents, the completeness of the new taxon provides an improved phylogenetic context for understanding the previously reported Gondwanan maniraptorans and their biogeographic history.

Theropoda Marsh, 1881

Maniraptora Gauthier, 1986

Dromaeosauridae Matthew and Brown, 1922

Unenlagiinae Bonaparte, 1999, *nov. comb.*

Buitreraptor gonzalezorum, gen. et sp. nov.

Holotype. MPC 245 (Museo Carlos Ameghino, Cipolletti, Río Negro Province, Argentina), consisting of a near-complete, articulated, adult skeleton (Figs 2 and 3a–h).

Referred specimen. MPC 238, an articulated partial skeleton comprising the hips, right hindlimb and sacrum (Fig. 3i, j).

Etymology. From Buitrera ('vulture roost'), the type locality, and *raptor*, robber; and *gonzalezorum*, in honour of the brothers Fabián and Jorge González, who discovered the holotype, for their dedicated participation in 'La Buitrera' fieldwork over the years.

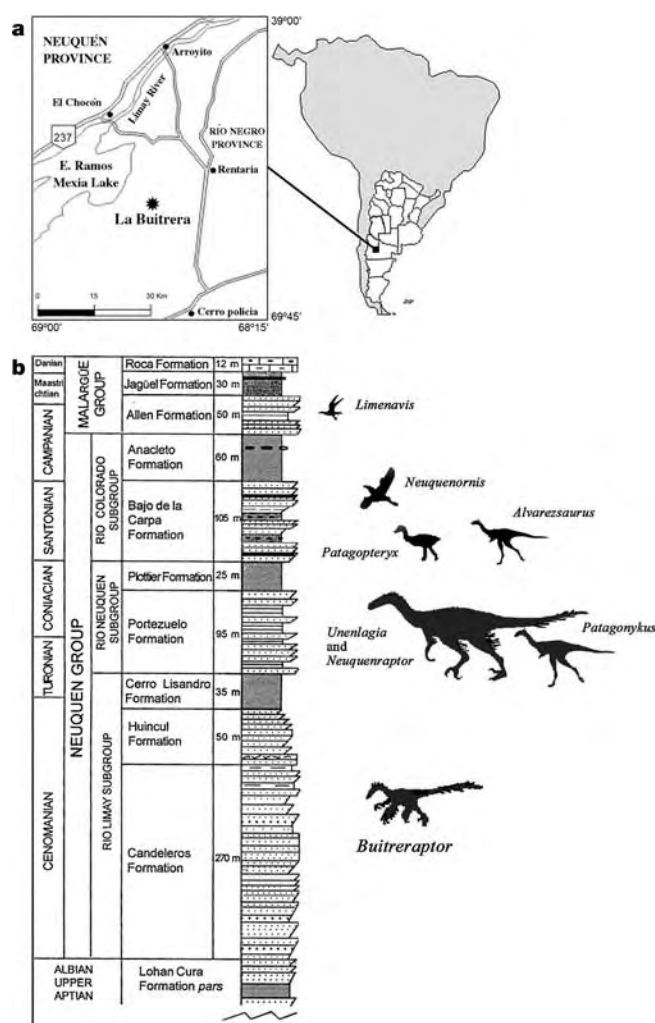


Figure 1 | Geographic and stratigraphic provenance of *Buitreraptor gonzalezorum*. **a**, Map of northwestern Patagonia showing the 'La Buitrera' locality between the villages of El Chocón and Cerro Policia. 'La Buitrera' is located about 80 km southwest of Cipolletti close to the south shore of Lake Ezequiel Ramos-Mexía. The approximate location of the map area is indicated on the South American continent. **b**, Stratigraphic column of the Neuquén Group (Upper Cretaceous) showing horizons of maniraptorans described from the Neuquén and Malargüe groups.

¹Department of Geology, The Field Museum, 1400 South Lake Shore Drive, Chicago, Illinois 60605, USA. ²Sección de Paleontología de Vertebrados, ³Laboratorio de Anatomía Comparada, Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia', Av. Ángel Gallardo 470, Buenos Aires (1405), Argentina. ⁴Fundación de Historia Natural 'Félix de Azara', CAECNA, Universidad Maimónides, V. Virasoro 732, Buenos Aires (1405), Argentina.

Locality and horizon. ‘La Buitrera’ facies of the Candeleros Formation (Cenomanian–Turonian)¹⁹, Río Negro Province, northwest Patagonia, Argentina (Fig. 1). The specimen was found in fluvial sandstones that yielded abundant skeletons of small vertebrates^{14–16} and dinosaurs^{20–21}.

Diagnosis. *Buitreraptor* differs from other theropods in the following unique combination of traits: skull long, exceeding femoral length by 25%; teeth small, unserrated, without root–crown constriction; quadrate with large lateral flange and pneumatic foramen; posterior cervical centra with ventrolateral ridge; furcula pneumatic; brevis shelf expanded and lobate, projects laterally from caudal end of ilium.

The holotype skeleton of *Buitreraptor* is almost complete, lacking only parts of the distal hindlimbs and forelimbs and the distal half of the tail, although many bones are scavenged. Complete fusion of neurocentral sutures indicates that this animal was probably an adult.

The skull (Fig. 2a, b) is long and low and considerably exceeds femoral length, in contrast to most other theropods. The large maxillary fenestra is separated from the antorbital fenestra by a narrow interfenestral bar (Fig. 2b, d). It is positioned near the middle of the skull as in some other paravians², and the antorbital fossa extends anterior to it. The nasals are flat and dorsally narrow, suggesting a long, thin muzzle as in *Velociraptor*. The orbital margins of the frontals flare laterally (Fig. 2a) and the triangular skull table indicates a high degree of stereoscopic vision. The body of the jugal is low, with a short, angled postorbital process, as in basal paravians. The postorbital bar is complete (Fig. 2b). Each quadrate is monostylic and bears a wide lateral process that defines the medial edge of an enlarged quadrate foramen (Fig. 2e) as in dromaeosaurids²². A

pneumatic foramen pierces the caudal face of the shaft as in a variety of theropod groups²³.

The labial surface of the dentary bears a deep subalveolar groove occupied by mental foramina (Fig. 2d), as in Troodontidae^{3,23}. A derived character shared with deinonychosaurs²² is the fusion of the interdermal plates. Atypically for dromaeosaurids, the widely spaced teeth are minute and lack serrations, but have unconstricted root–crown transitions (Fig. 2b–d).

The cervicals bear low neural spines and small epiphyses. Posterior cervical centra bear carotid processes (Fig. 3a) as in many paravians^{1,23}. Low ridges form the ventrolateral corners of the last cervical centrum and terminate caudally as small tubers, a feature unique to *Buitreraptor*. Hypapophyses are developed on the anterior trunk vertebrae as in most maniraptorans. Dorsal neural spines are tall and rectangular as in dromaeosaurids¹ (Fig. 3b). Another dromaeosaurid synapomorphy is the peduncular shape of the parapophyses on the trunk vertebrae¹. *Buitreraptor* has a long tail, which, although incomplete in both specimens, is comparable to that of *Archaeopteryx*, *Jeholornis* and *Microraptor*. The transition between proximal and distal caudals occurs between caudals 7 and 8. Distal caudals do not bear a neural spine, but unlike Laurasian dromaeosaurids, the prezygapophyses are tapered and only overlap half of the preceding vertebra (Fig. 3c). As in other paravians, the dorsoventrally compressed chevrons are bifid at both ends.

Buitreraptor has a broad, robust furcula with curved rami and a low ridge instead of a hypocleideum (Fig. 3d). The furcula is hollow with trabeculae spanning its interior. The scapular blade is long and straplike, with a pronounced curvature near the glenoid, which is directed laterally as in paravians^{2,5}. The enlarged coracoids display a

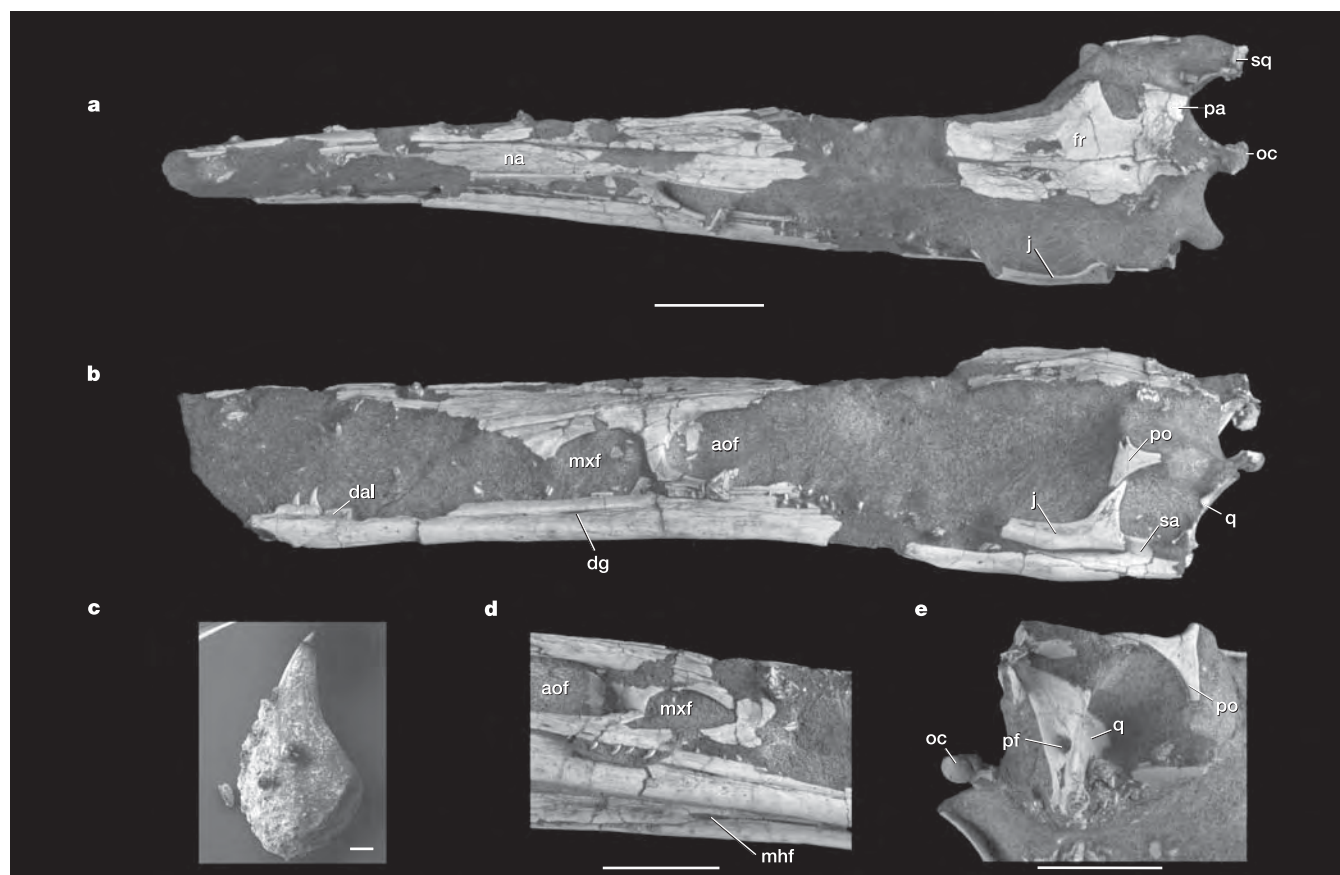


Figure 2 | *Buitreraptor gonzalezorum* MPCA 245, holotype. **a**, Skull in dorsal view. **b**, Skull in left lateral view. **c**, Isolated tooth crown. **d**, Close-up of right maxillary fenestra. **e**, Close-up of right quadrate. Scale bars, 20 mm (**a**, **b**, **d**, **e**) and 1 mm (**c**). Abbreviations: aof, antorbital fenestra; dal, dentary

alveolus; dg, dentary groove; fr, frontal; j, jugal; mh, mylohyoid foramen; mx, maxillary fenestra; na, nasal; oc, occipital condyle; pa, parietal; pf, pneumatic foramen; po, postorbital; q, quadrate; sa, surangular; sq, squamosal.

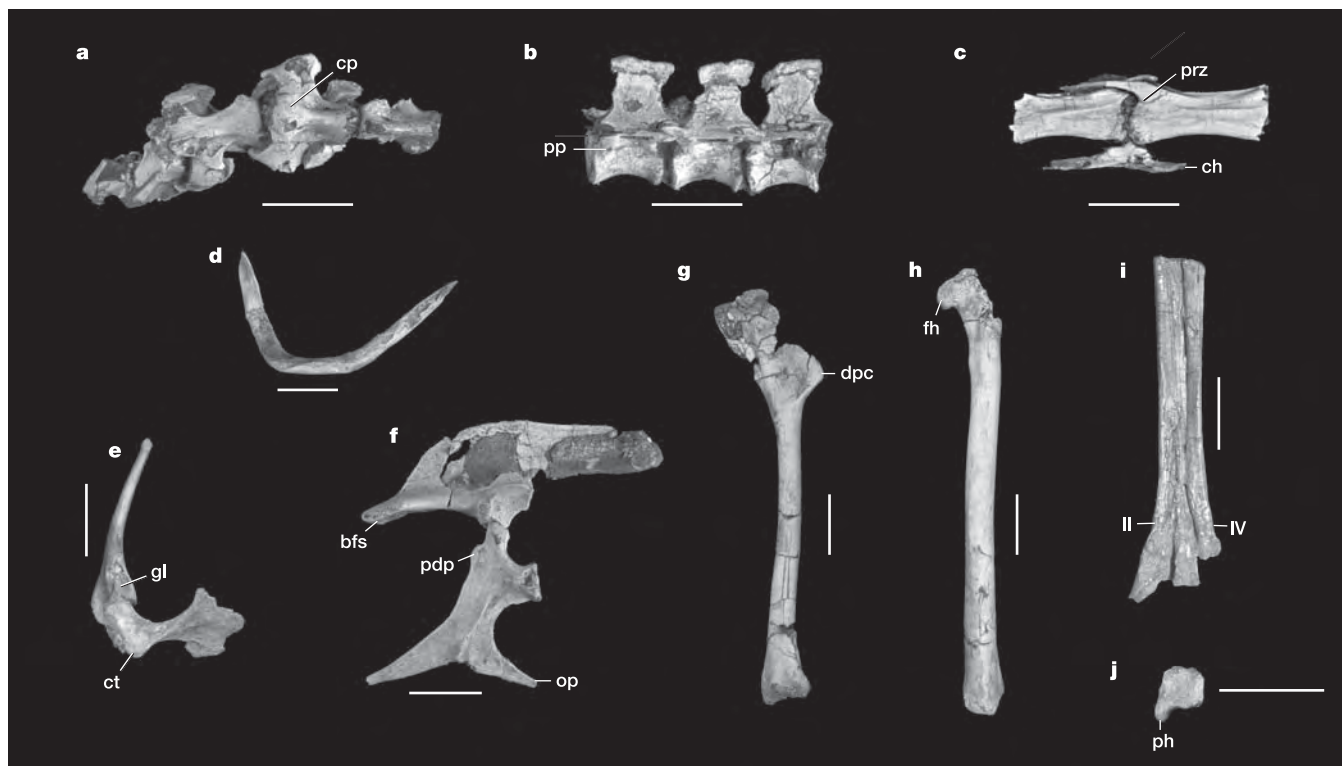


Figure 3 | *Buitreraptor gonzalezorum*. **a–h**, MPCA 245; **i, j**, MPCA 238. **a**, Cervicodorsal vertebrae, ventral view. **b, c**, Midtrunk vertebrae (**b**) and distal caudal vertebrae (**c**), lateral views. **d, e**, Furcula (**d**) and left scapulocoracoid (**e**), caudal views. **f, g**, Right humerus (**g**) and ilium and ischium (**f**), lateral views. **h**, Left femur, anterior view. **i**, Right metapodium,

plantar view. **j**, Right phalanx II-2, proximal view. Scale bars, 20 mm. Abbreviations: bfs, brevis fossa; ch, chevron; cp, carotid process; ct, coracoid tuber; dpc, deltopectoral crest; fh, femoral head; gl, glenoid; op, obturator process; pdp, posterodorsal process; ph, flexor heel; pp, parapophysis; prz, prezygapophysis.

perpendicular flexure between the glenoid portion and the expanded ventral blade (Fig. 3e) as in *Sinornithosaurus*². A prominent biceps tuber occupies the edge of this flexure and is flanked by a wide subglenoid fossa as in dromaeosaurids and basal avians.

The humerus is 30% longer than the scapula, an unusual ratio only matched or surpassed in avians and some dromaeosaurids from Liaoning^{1,2}. It bears a large deltopectoral crest and a straight-edged internal tuberosity (Fig. 3g). The ulna is bowed posteriorly and has a subdivided proximal articulation, as in other paravians. Parts of three metacarpals are preserved, with metacarpal I shorter than the other metacarpals.

The pelvis is better preserved on the referred specimen. A distinct supra-acetabular crest is present on the ilium (Fig. 3f) as in *Unenlagia*⁵ and *Rahonavis*¹⁰. The brevis shelf is lobate and laterally directed, and extends beyond the vertical lamina of the ilium. It is reminiscent of the ilium of *Microraptor*¹ and to a lesser degree *Unenlagia*^{5,7,12}. The pubis is oriented vertically, but the pubic shaft has a strong anteriorly convex curvature as in dromaeosaurids, including *Unenlagia paynemili*⁷. The ischium (Fig. 3f) bears a proximally positioned dorsal process as in many paravians, and has a distally placed obturator process that is long and acutely pointed as in *Rahonavis*¹⁰, *Microraptor*¹ and *Sinornithosaurus*². A sharp, longitudinal crest subdivides the lateral face of the shaft as in the latter taxon.

The femur (Fig. 3h) has a ventrolaterally directed head without a distinct neck, and the shaft is bowed anteriorly. It bears a well-developed lateral ridge but barely a trace of the fourth trochanter. The tibia is longer than the femur and the slender fibula reaches the ankle. The foot of the referred specimen is arctometatarsal with a wedge-shaped third metatarsal that remains visible, but recessed, along the plantar and extensor faces (Fig. 3i), as in *Neuquenraptor*⁶. Metatarsals II and IV are symmetric around metatarsal III, unlike the condition in

troodontids^{3,23}, and metatarsals II and III terminate in ginglymoid articulations. Phalanx II-2 bears a proximoventral flexor heel and an extensive distal articulation as in deinonychosaurs. The flexor heel is weak and offset medially (Fig. 3j) as in *Rahonavis*¹⁰, *Neuquenraptor*⁶, *Microraptor*¹ and some primitive troodontids³. A natural mould of a hypertrophied digit II ungual was preserved with the referred specimen.

Phylogenetic analysis (Supplementary Information) posits *Buitreraptor* within Dromaeosauridae (Fig. 4a). Characters uniting *Buitreraptor* with dromaeosaurids include the lateral process of the quadrate (Fig. 2e), unconstricted teeth (Fig. 2c), stalked trunk parapophyses (Fig. 3b), bifid chevrons and a ginglymoid distal articulation on metatarsal II. This analysis unites the purported primitive avialan *Rahonavis* with Dromaeosauridae. In part because of its forearm proportions (Supplementary Information), *Rahonavis* has been interpreted as a flying bird¹⁰, but it shares derived pelvic and hindlimb features with *Buitreraptor* and *Unenlagia*¹². Although proportionally shorter than in *Rahonavis*, the forelimb bones are also elongate in *Buitreraptor* and other basal dromaeosaurids^{1,2}. According to our phylogenetic results, the particularly elongate forelimbs of *Rahonavis* and inferred potential flight capability originated independently of Aves (Supplementary Information).

Two other basal dromaeosaurids from Argentina, *Neuquenraptor*⁶ and *Unenlagia*^{5,7}, are based on fragmentary specimens from the Portezuelo Formation, which is younger than the Candeleros Formation (Fig. 1). Their incompleteness and poor degree of skeletal overlap obscure their exact phylogenetic positions^{5–7,12}. Co-occurrence within the same formation and geographic area, and general size considerations, suggest that the taxa could be congeneric. Near-identical dimensions and proportions of the modified phalanges in the raptorial second toe of *Neuquenraptor* and material assigned to *Unenlagia paynemili*⁷ further support the possible

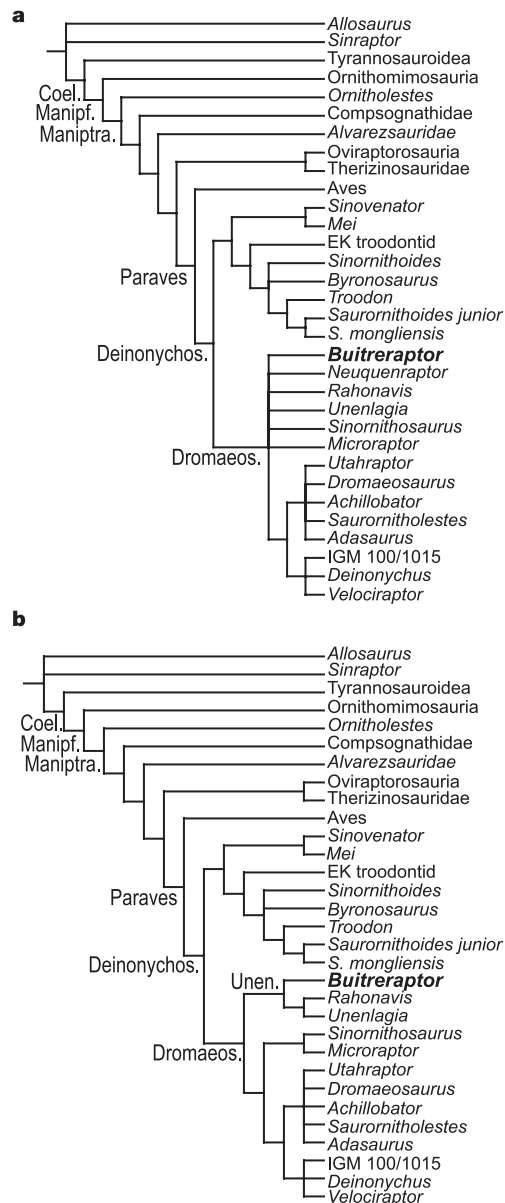


Figure 4 | Relationships of *Buitreraptor gonzalezorum* within Coelurosauria. **a**, Strict consensus of 504 trees depicting phylogenetic relationships of *Buitreraptor gonzalezorum* within Coelurosauria, based on a parsimony analysis^{28,29} of 236 characters in 58 genera. Clades basal to Paraves are represented by suprageneric branches but were sampled at specific level. Suprageneric taxonomy follows ref. 30 except for Therizinosauridae and Unenlagiinae. See Supplementary Information for character list, data matrix, analysis protocol, node diagnoses and discussion. **b**, Strict consensus of 36 trees with *Neuquenraptor* treated as a junior synonym of *Unenlagia*, yielding a monophyletic Unenlagiinae. Abbreviations: Coel., Coelurosauria; Deinonychos., Deinonychosauria; Dromaeos., Dromaeosauridae; Manipf., Maniraptoriformes; Maniptra., Maniraptora; Unen., Unenlagiinae.

synonymy. A revised phylogeny with *Unenlagia* and *Neuquenraptor* combined unites *Buitreraptor*, *Unenlagia* and *Rahonavis* within a Gondwanan lineage of dromaeosaurids, the Unenlagiinae²⁴, distinct from the better-known Laurasian diversity of this clade (Fig. 4b). A large unnamed deinonychosaurian from the uppermost Cretaceous of Argentina²⁵ shares derived dental traits with *Buitreraptor*, indicating that unenlagiines may represent a long-lived lineage with wide ranging body sizes, thus mirroring some aspects of the evolutionary history of Laurasian dromaeosaurids³. Moreover, the morphological disparity of unenlagiines such as the longirostrine skull of

Buitreraptor and the wing-like forelimbs of *Rahonavis*, exceeds that observed in the better-known Laurasian radiation.

The Gondwanan occurrences of taxa such as the alvarezsaurid *Patagonykus*⁴ and *Unenlagia*⁵, coupled with the basal positions that they occupy within their respective subclades, have been interpreted as evidence for a vicariant distribution of maniraptoran lineages after the Late Jurassic breakup between northern and southern landmasses⁶. By providing evidence for uniting Gondwanan dromaeosaurids into an endemic lineage, *Buitreraptor* corroborates this vicariant hypothesis and also underscores the profound separation between Laurasian and Gondwanan faunas, which evolved in parallel for most of the Cretaceous period²⁶. This biogeographic pattern is further bolstered by the faunal elements associated with *Buitreraptor*, such as dryolestoid mammals and eilenodontine spheonodontians¹⁵, which represent relicts of more global Jurassic distributions. A Gondwanan distribution after Pangaeian break-up implies an unenlagiine origin in the Middle or Late Jurassic period, thus providing indirect evidence for the radiation of the principal maniraptoran lineages before the appearance of birds in the fossil record²⁷.

Received 6 May; accepted 5 July 2005.

- Hwang, S. H., Norell, M. A., Ji, Q. & Gao, K.-Q. New specimens of *Microraptor zhaoianus* (Theropoda: Dromaeosauridae) from Northeastern China. *Am. Mus. Novit.* **3381**, 1–44 (2002).
- Xu, X., Wang, X.-L. & Wu, X.-C. A dromaeosaurid dinosaur with filamentous integument from the Yixian Formation of China. *Nature* **401**, 262–266 (1999).
- Xu, X., Norell, M. A., Wang, X.-L., Makovicky, P. J. & Wu, X.-C. A basal troodontid from the Early Cretaceous of China. *Nature* **415**, 780–784 (2002).
- Novas, F. E. Alvarezsauridae, Late Cretaceous maniraptorans from Patagonia and Mongolia. *Queensland Mus. Mem.* **39**, 675–702 (1996).
- Novas, F. E. & Puerta, P. F. New evidence concerning avian origins from the Late Cretaceous of Patagonia. *Nature* **387**, 390–392 (1997).
- Novas, F. E. & Pol, D. New evidence on deinonychosaurian dinosaurs from the Late Cretaceous of Patagonia. *Nature* **3285**, 858–861 (2005).
- Calvo, J. O., Porfiri, J. D. & Kellner, A. W. On a new maniraptoran dinosaur (Theropoda) from the Upper Cretaceous of Neuquén, Patagonia, Argentina. *Arq. Mus. Nacional. Rio de Janeiro* **62**, 549–566 (2004).
- Novas, F. E. & Agnolín, F. L. *Unquillosaurus ceibali* Powell, a giant maniraptoran (Dinosauria, Theropoda) from the Late Cretaceous of Argentina. *Rev. Mus. Argentino Cienc. Nat.* **NS** **6**, 61–66 (2004).
- Rauhut, O. W. & Werner, C. First record of the family Dromaeosauridae (Dinosauria: Theropoda) in the Cretaceous of Gondwana (Wadi Milk Formation, Northern Sudan). *Paläont. Z.* **69**, 475–489 (1995).
- Forster, C. A., Sampson, S. D., Chiappe, L. M. & Krause, D. The theropod ancestry of birds: new evidence from the Late Cretaceous of Madagascar. *Science* **279**, 1915–1919 (1998).
- Clark, J. M., Norell, M. A. & Makovicky, P. J. in *Mesozoic Birds: Above the Heads of Dinosaurs* (eds Chiappe, L. M. L. & Witmer, L.) 31–64 (Univ. of California Press, Berkeley, 2002).
- Novas, F. E. in *Feathered Dragons: the Origin of Birds and Flight* (eds Currie, P. J., Koppelhus, E., Shugar, M. & Wright, J. A.) 150–166 (Indiana Univ. Press, Bloomington, 2004).
- Sereno, P. C. in *New Perspectives on the Origin and Evolution of Birds: Proceedings of the International Symposium in Honor of John H. Ostrom* (eds Gauthier, J. A. & Gall, L. F.) 69–98 (Peabody Mus. Nat. Hist., Yale Univ., New Haven, 2001).
- Apesteguía, S., de Valais, S., González, J. A., Gallina, P. A. & Agnolín, F. L. The tetrapod fauna of 'La Buitrera', new locality from the basal Late Cretaceous of North Patagonia, Argentina. *J. Vertebr. Paleontol.* **21**, 29A (2001).
- Apesteguía, S. & Novas, F. E. Large Cretaceous spheonodontian from Patagonia provides insight into lepidosaur evolution in Gondwana. *Nature* **425**, 609–612 (2003).
- Carignano, A. P., Pol, D., Apesteguía, S. & Novas, F. E. La diversidad de cocodrilos de 'La Buitrera' (Fm. Candeleros), provincia de Río Negro. *Ameghiniana* **39**, 7R (2002).
- Coria, R. & Salgado, L. A new giant carnivorous dinosaur from the Cretaceous of Patagonia. *Nature* **377**, 224–226 (1995).
- Calvo, J. O. & Salgado, L. *Rebbachisaurus tessonei* sp. nov., a new Sauropoda from the Albian–Cenomanian of Argentina: new evidence on the origin of the Diplodocidae. *Nat. Sci. Mus. Monogr.* **15**, 13–45 (1995).
- Corbella, H., Novas, F. E., Apesteguía, S. & Leanza, H. A. First fission-track age for the dinosaur-bearing Neuquén Group (Upper Cretaceous), Neuquén basin, Argentina. *Rev. Mus. Argentino Cienc. Nat.* **NS** **6**, 227–232 (2004).
- de Valais, S. & Apesteguía, S. Dientes asignables a *Giganotosaurus* (Carcharodontosauria, Theropoda) provenientes de 'La Buitrera', Fm. Candeleros, Provincia de Río Negro. *Ameghiniana* **38**, 9R (2001).
- Gallina, P. A., Apesteguía, S. & Novas, F. E. ¿Un elefante bajo la alfombra? Los

- Rebbachisauridae (Sauropoda, Diplodocimorpha) del Cretácico de Gondwana. Nuevas evidencias en 'La Buitrera' (Fm. Candeleros), provincia de Río Negro. *Ameghiniana* **39**, 10R (2002).
22. Currie, P. J. New information on the anatomy and relationships of *Dromaeosaurus albertensis* (Dinosauria: Theropoda). *J. Vertebr. Paleont.* **15**, 576–591 (1995).
 23. Makovicky, P. J., Norell, M. A., Clark, J. M. & Rowe, T. Osteology and relationships of *Byronosaurus jaffei* (Theropoda: Troodontidae). *Am. Mus. Novit.* **3402**, 1–32 (2003).
 24. Bonaparte, J. F. Tetrapod faunas from South America and India: A palaeobiogeographic interpretation. *Proc. Indian Nat. Sci. Acad.* **65**, 427–437 (1999).
 25. Novas, F. E., Canale, J. I. & Isasi, M. Giant deinonychosaurian theropod from the Late Cretaceous of Patagonia. *J. Vertebr. Paleont.* **24** (Suppl.), 98A (2004).
 26. Bonaparte, J. F. History of the terrestrial Cretaceous vertebrates of Gondwana. *Actas IV Congr. Argentino Paleontol. Bioestrat.* **4**, 63–95 (1986).
 27. Brochu, C. A. & Norell, M. A. Temporal congruence and the origin of birds. *J. Vertebr. Paleont.* **20**, 197–200 (2000).
 28. Goloboff, P. A. NONA v.2.0. (published by the author, S.M. de Tucumán, Argentina, 1993).
 29. Nixon, K. Winclada (BETA) v.0.9.9. (published by the author, Ithaca, New York, 2000).
 30. Sereno, P. C. A rationale for phylogenetic definitions, with application to the

higher-level taxonomy of Dinosauria. *N. Jb. Geol. Paläont. Abh.* **210**, 41–83 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Avelás and Salinas families, C. Muñoz, L. Fernández and the 2004 La Buitrera field crew for their help in the field; J. A. González, P. Chiarelli, M. Brown, J. Holstein, L. Herzog, A. Shinya, C. Van Beek and D. Wagner for preparing the holotype; J. Weinstein for photography, M. H. Donnelly for retouching Fig. 1; J. González for the reconstruction of *Buitreraptor*; S. de Valais for help and discussion during the early stages of this work; D. Pol for reading the manuscript and providing advice; and F. E. Novas and R. A. Coria for providing access to materials in their care. Fieldwork and preparation was facilitated by the Agencia Cultura of Río Negro Province and supported by The Jurassic Foundation (S.A.) and NASA (P.J.M.), and research was supported by the NSF (P.J.M.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.J.M. (pmakovicky@fieldmuseum.org) or S.A. (paleoninja@yahoo.com.ar).

LETTERS

Further evidence for small-bodied hominins from the Late Pleistocene of Flores, Indonesia

M. J. Morwood^{1*}, P. Brown^{1*}, Jatmiko², T. Sutikna², E. Wahyu Saptomo², K. E. Westaway³, Rokus Awe Due², R. G. Roberts³, T. Maeda¹, S. Wasisto² & T. Djubiantono²

Homo floresiensis was recovered from Late Pleistocene deposits on the island of Flores in eastern Indonesia, but has the stature, limb proportions and endocranial volume of African Pliocene *Australopithecus*¹. The holotype of the species (LB1), excavated in 2003 from Liang Bua, consisted of a partial skeleton minus the arms. Here we describe additional *H. floresiensis* remains excavated from the cave in 2004. These include arm bones belonging to the holotype skeleton, a second adult mandible, and postcranial material from other individuals. We can now reconstruct the body proportions of *H. floresiensis* with some certainty. The finds further demonstrate that LB1 is not just an aberrant or pathological individual, but is representative of a long-term population that was present during the interval 95–74 to 12 thousand years ago. The excavation also yielded more evidence for the depositional history of the cave and for the behavioural capabilities of *H. floresiensis*, including the butchery of *Stegodon* and use of fire.

In 2003, archaeological excavations at Liang Bua, in western Flores, yielded a partial hominin skeleton (LB1) that proved to be about 18 thousand years (kyr) in age^{1,2}. The skeleton is the holotype for a new species, *Homo floresiensis*, which we argued was the result of endemic dwarfing of an earlier *H. erectus* population¹.

In 2004, we continued the excavation in Sector VII (from which the holotype skeleton was recovered) to a depth of 11 m without encountering bedrock. The adjoining Sector XI, located immediately to the south, was also excavated to recover further hominin remains, including additional elements of LB1, and to obtain more evidence for their stratigraphic and archaeological contexts (see Supplementary Information). Both sectors were excavated by stratigraphic layers as well as by 10-cm-deep spits.

The stratigraphic section of the combined Sector VII/XI trench indicates a consistent pattern of sediment deposition in this part of the cave during the Late Pleistocene (Fig. 1). At the northern, lower-lying end of the trench, layers of clayey silt (suspension deposits) accumulated in pools of standing water, with intrusions of gravel reflecting times of strong water flow from higher ground to the south. Elements of the LB1 skeleton, with some parts still articulated¹, were recovered in both Sectors VII and XI, from a clayey silt (Layer R) accumulated on the lee side of a conglomerate deposit (Layer S). This suggests that natural processes were responsible for the rapid burial of the body by fine-grained sediments in a standing body of water. This pattern of sediment deposition changed about 12–11 kyr ago, when the lower sections of the cave floor were filled with thick, 'white' tuffaceous silts, another phase of suspension deposition, as evident in Sectors IV, VII and XI (see Supplementary Information). The overlying Holocene deposits were generally laid down horizontally by moving water.

A tiny hominin radius from Spit 42 in Sector XI was found immediately below the 'white' tuffaceous silts, and 40 cm above a calibrated radiocarbon age of 13.1 kyr from Sector VII (ANUA-27115; ref. 2). The inferred age of 12 kyr for the radius provides the youngest evidence for *H. floresiensis* and *Stegodon* at Liang Bua. In Sector VII, thermoluminescence dating of quartz grains from depths of 935 and 862 cm yielded maximum ages for sediment deposition of 55 ± 8 kyr (sample LBS7-45a) (mean $\pm$ total uncertainty at the 68% confidence interval) and 41 ± 10 kyr (sample LBS7-46a) respectively, using the light-sensitive red thermoluminescence signal (see Supplementary Information). Although at a greater depth, the artefact- and bone-bearing deposits in this part of the site are not as old as those in the basal levels of Sector IV, where the earliest evidence for *H. floresiensis*, *Stegodon* and stone artefacts occurs at 95–74 kyr (ref. 2).

The uppermost 3.7 m of deposit in Sector XI yielded a complete Holocene sequence in which all hominin remains are of modern humans (*H. sapiens*). In contrast, none of the hominin remains recovered from the Pleistocene deposits are of modern humans. Instead, all diagnostic hominin elements are consistent with *H. floresiensis* size and morphology (Table 1). In total, there are now at least nine such individuals represented in the excavated finds from Sectors IV, VII and XI. This paper is concerned mainly with another mandible (LB6/1) and tibia (LB8), as well as the right humerus and ulna of LB1—all excavated from Sector XI in 2004.

During the Late Pleistocene, the southern, topographically higher area in the Sector VII/XI trench comprised dry, relatively level ground adjacent to a pool against the east wall of the cave. High densities of stone cores, flaking debris, retouched tools, anvils and faunal remains were found there, indicating that the area was a focus for a range of hominin activities and that *H. floresiensis* was capable of complex behaviour and cognition (see Supplementary Information). However, no pigments or symbolic items were found, in contrast with Holocene occupation levels.

Faunal remains included those of *Stegodon* (with cut marks evident on some bones), Komodo dragon, rat and bat, as well as the hominin mandible, scapula, radius and ulna from Spit 51 that were found on a well-defined occupation floor. Use of fire by hominins is indicated by charred bone and clusters of reddened and fire-cracked rocks. These include a cluster of three burnt, water-rolled, volcanic pebbles from Spit 84 (840 ± 5 cm depth) in Sector VII, and a circular arrangement of five similarly burnt pebbles from Spit 43 (435 ± 5 cm depth) in Sector XI.

With regards to the hominin remains, the adult mandible (LB6/1) from Spit 51 is undistorted, but has post-mortem damage to the left coronoid process and condyle. Most of the right ramus and the central and left lateral incisors were also lost post-mortem (Fig. 2).

¹Archaeology and Palaeoanthropology, School of Human and Environmental Studies, University of New England, Armidale, New South Wales 2351, Australia. ²Indonesian Centre for Archaeology, Jl. Raya Condok Pejaten No. 4, Jakarta 12001, Indonesia. ³GeoQuEST Research Centre, School of Earth and Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia.

*These authors contributed equally to this work.

The mandibular dental arch is narrow anteriorly and long relative to its breadth. The axis of P_3 – M_3 is straight, forming a more V-shaped tooth row than in the holotype, and in lateral view the inferior border of the mandible is not as arched as LB1. All permanent teeth had erupted and show evidence of occlusal wear, but less so than in LB1, and the Curve of Spee is not as well developed. The ramus root inserts on the corpus above the lateral prominence, and in lateral aspect obscures the distal M_3 . The ramus is broadest inferiorly, slopes slightly posteriorly and is thickened medio-laterally, and the extramolar sulcus is broader than in LB1, with a moderate lateral prominence. Bilaterally, there are multiple mental foramina in LB6/1, which is a common feature in Asian *H. erectus*^{3,4}.

The anterior portion of the corpus is without a chin (no mental protuberance, mental tubercles or incurvatio mandibularis), but is not as rounded and bulbous as LB1. In the posterior symphyseal region, the alveolar planum inclines postero-inferiorly; there is a

moderate superior torus; the diagastric fossa is deep and broad; the inferior transverse torus is low and rounded, rather than shelf-like; and there is a strong posterior angulation of the symphyseal axis. D211 *H. erectus* from Dmanisi, Georgia, has similar symphyseal thickening and internal transverse tori, but differs in having a small mental protuberance⁵.

With the exception of P_3 , the size and morphology of the LB6/1 mandibular teeth follow the pattern in *H. erectus* and *H. sapiens*, and are consistent with the previously described attributes of *H. floresiensis*¹: the molar size sequence is $M1 \geq M2 > M3$. As with LB1 and LB2 (ref. 1), the P_3 s are mesio-distally elongated, and had a prominent protoconid and broad talonid when unworn. Both P_3 s also have bifurcated roots and the P_4 s have a mesiodistally compressed, broad Tomes' root. Mandibular P_3 s and P_4 s with similar crown and root morphology have been recorded for *Australopithecus* and early *Homo*^{6,7}. Some Indonesian *H. erectus* mandibular premolars also have bifurcated or Tomes' roots⁴, as does Dmanisi D2600

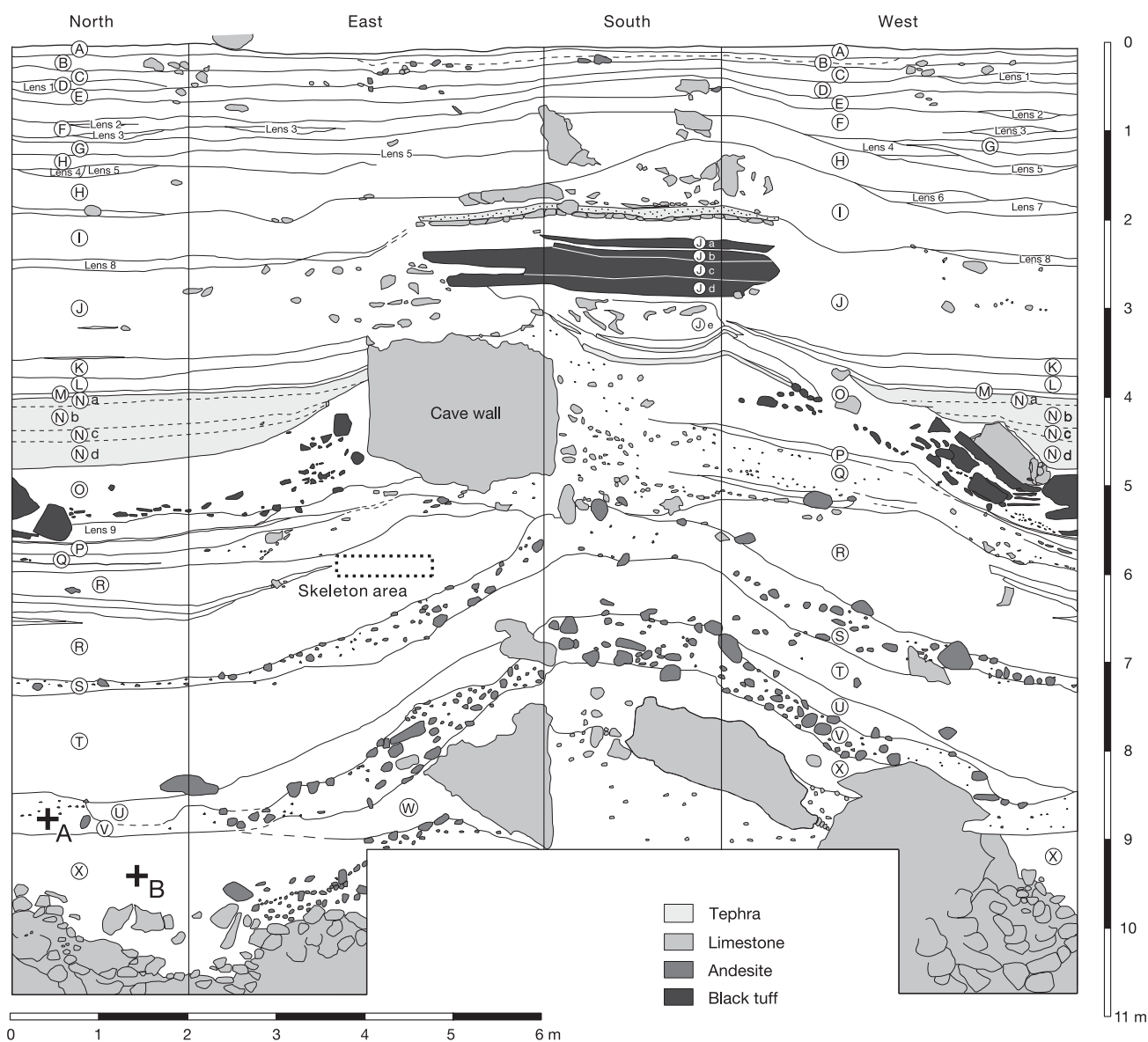


Figure 1 | Stratigraphy in the Sector VII/XI trench, Liang Bua. Layers A–M are brownish clayey silts and sandy silts. N(a–d) comprise ‘white’ tuffaceous silts dated 12–11 kyr old. Evidence for *H. floresiensis* and *Stegodon* occurs up to the base of these. Layers O–R, T, U, W and X are brownish clayey silts and silts varying in texture, colour and proportions of rock. The holotype of

H. floresiensis occurred in Layer R, a clayey silt. Layers S and V are conglomerates with clayey silt matrix, from stronger water flow. Crosses labelled A and B indicate the locations of thermoluminescence samples LBS7-46a and LBS7-45a, respectively.

Table 1 | Hominin remains excavated from the Late Pleistocene levels of Sector XI, Liang Bua, in 2004

Spit no.	Bone	Description
42	Radius: child, left (LB4/1)	Proximal epiphyses unfused, articular surfaces not recovered, and distal quarter of shaft incomplete. Maximum length of fragment 101 mm.
43	Tibia: child, right (LB4/2)	Distal and proximal epiphyses unfused and articular surfaces not recovered. Maximum length 117 mm. Distal end recovered from Spit 44.
46	Cervical vertebra C1 (LB5/1) Metacarpal: adult (LB5/2)	Incomplete, represented by two fragments. Proximal end broken, length 58 mm.
51	Mandible: adult (LB6/1) Radius: adult, right (LB6/2) Ulna: adult, left (LB6/3) Scapula: adult, right (LB6/4) Metatarsal (LB6/5) Phalanx: 1st of foot (LB6/6) Phalanx: 3rd of hand (LB6/7)	With incomplete left ramus and right coronoid process and condyle. Originally with a fracture through the corpus between right P ₃ and P ₄ and left M ₁ and M ₂ . Subsequently broken at the symphysis when removed from the Centre for Archaeology in Jakarta. This has altered the original arch dimensions, occlusion and the morphology of the symphysis. At the same time, cut marks, fill and glue altered the morphology of the lateral corpus and ramus. Complete. It has an angulated, healed fracture in the distal third with compensatory remodelling and extensive callus development. The forearm would have been bowed and distorted and movement of the hand restricted. Maximum length 157 mm. Proximal shaft. Maximum length 137 mm. With incomplete superior border, medial spine, inferior angle and coracoid process. Maximum breadth of glenoid cavity 25.1 mm, length of the auxiliary border 82 mm. Articular surfaces not preserved. Articular surfaces not preserved. Complete. Maximum length 10 mm.
52	Phalanx: 1st of hand (LB6/8) Phalanx: 2nd of hand (LB6/9) Phalanx: 2nd of hand (LB6/10) Phalanx: 1st of hand (LB6/11) Phalanx: 3rd of hand (LB6/12) Phalanx: 1st of foot (LB6/13)	Complete. Maximum length 30.5 mm. Complete. Maximum length 16 mm. Distal end not preserved. Complete. Maximum length 10.5 mm. Complete. Maximum length 12.5 mm. Complete. Maximum length 16 mm.
53	Incisor: mandibular I ₁ (LB6/14)	
56B	Phalanx: 1st of hand (LB7)	Proximal end not preserved.
58A*	Humerus: adult, right (LB1) Ulna: adult, right (LB1) Fibula: adult, left (LB1) Tibia: adult, right (LB8) Ulna: adult, left (LB1)	Lateral epicondyle and capitulum are incomplete, and the greater and lesser tubercles are not preserved. Maximum length 243 mm. Distal end of shaft and head not preserved. Estimated maximum length 205 mm. Complete but with fracture through distal end of shaft. Maximum length 226 mm. This pairs with the right fibula of LB1 recovered from Sector VII in 2003 and is the same length. Medial condyle incomplete and medial malleolus not preserved. Estimated maximum length 216 mm. This tibia duplicates that found with the skeleton and is smaller. It is therefore from another individual. Reassembled when removed from the Centre for Archaeology in Jakarta. Now has altered morphology and adhering glue. Both epiphyses missing. 167 × 16.5 mm. (Note: from baulk collapse.)
65B	Femur (LB9)	Fragment. Shaft split longitudinally. 91 × 17 mm.

* The northern part of Spit 58 in Sector XI (designated 58A) was part of Layer R, a clayey silt deposited in a body of standing water. In Sector VII, Layer R contained the partial skeleton (LB1).



Figure 2 | Occlusal and lateral views of LB1 and LB6/1 *H. floresiensis* mandibles. LB1 mandible (left) from Spit 59, Sector VII, and LB6/1 mandible (right) from Spit 51, Sector XI. Scale bar, 1 cm.



Figure 3 | LB8 right tibia morphology. Posterior and anterior views of the LB8 right tibia, from Spit 58A, and anterior view of an African female Pygmy tibia scaled to the same length. Scale bar, 1 cm.

(ref. 8). Dmanisi D2735 has P_3 s with a similar crown shape to LB6/1, but the associated root morphology has not been described.

The LB6/1 mandible has smaller tooth and corpus dimensions than that of LB1, and is probably from a slightly smaller adult^{1,9} (see Supplementary Information). It also has a more V-shaped tooth row and is thicker in the posterior corpus (see Supplementary Information). Apart from such evidence of individual variation, LB6/1 has numerous detailed similarities with LB1 in overall size, morphology and symphysis, but is about 15 kyr in age—that is, 3 kyr younger than LB1. The second mandible therefore provides another line of evidence that LB1 was not aberrant, but is instead representative of a long-term, morphologically unique, small-bodied population with a configuration of features never recorded in normal or pathological

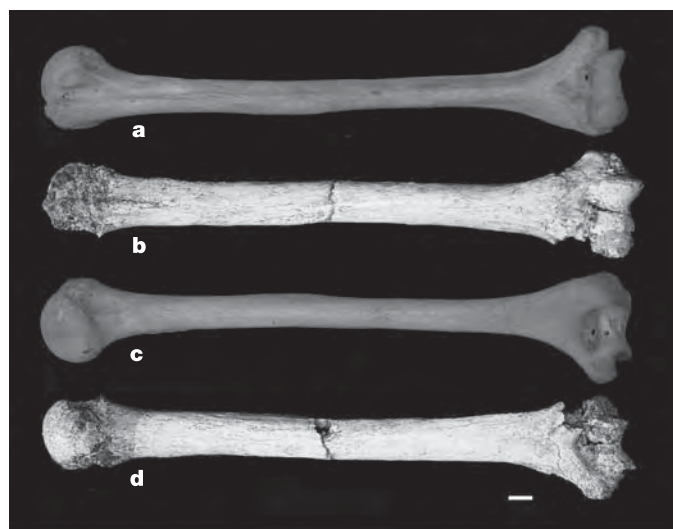


Figure 4 | LB1 right humerus morphology. Posterior (b) and anterior (d) views of the right humerus of LB1 *H. floresiensis*, with a small-bodied modern human (African Pygmy) humerus scaled to the same length (a, c). Scale bar, 1 cm.



Figure 5 | LB1 right ulna morphology. Lateral (a), posterior (b) and anterior (c) views of the right ulna of LB1 *H. floresiensis*, and anterior view (d) of a small-bodied modern human (African Pygmy) ulna scaled to the same length. Scale bar, 1 cm.

H. sapiens^{10–12}. In fact, morphologically and metrically, the LB1 and LB6/1 mandibles are more similar to those of early *H. erectus* in east Africa or Georgia than those in Java or China, whereas the symphyses most resemble those of Plio-Pleistocene hominins, such as LH4 *A. afarensis*¹³, and to a lesser extent KNM WT-15000 (ref. 14) and D211 (ref. 5) *H. erectus* (see Supplementary Information).

An extremely small adult right tibia (LB8), with fused distal and proximal epiphyses, was recovered from Spit 58A, in the same layer as LB1 (Fig. 3 and Supplementary Information). There is damage to the medial condyle and the medial malleolus is not preserved. In common with the holotype of *H. floresiensis*, the shaft is laterally concave and oval in cross-section (mid-shaft 194 mm²)—similar to *Pan*—without a sharp anterior border, and relatively thickened medio-laterally in the distal half compared with *H. sapiens*. Estimated maximum length of this tibia is 216 mm (see Supplementary Information), compared with 235 mm for LB1 (ref. 1). The relationship between midshaft circumference (56 mm) and the length of the tibia is in the *Pongo* and *Pan* range of variation, and distinct from *Homo*¹⁵.

Stature estimates for small-bodied hominins, with relative limb proportions and bone lengths similar to AL 281-1, are most appropriately estimated with coefficients derived from small-bodied modern humans, particularly African Pygmies^{16,17}. On the basis of Pygmy male coefficients¹⁶, the LB8 tibia indicates a stature of 109 cm

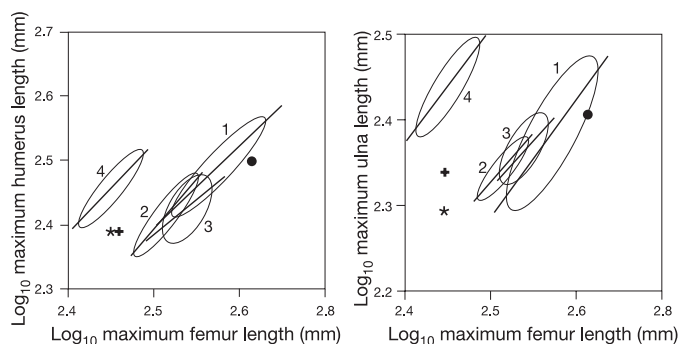


Figure 6 | Limb proportions in *H. floresiensis*. Relationship of humerus and ulna length in LB1 (star) to femur length, compared with those in KNM WT-15000 *H. erectus* (filled circle)⁴, AL 288-1 *A. afarensis* (cross)^{20,24}, modern humans (numbers 1–3) and *Pan paniscus* (number 4). The modern human comparative series is from the robust global sample used previously¹, with the addition of African Pygmies (2) and Andaman Islanders (3). Least squares regression lines and 95% confidence ellipses are shown.

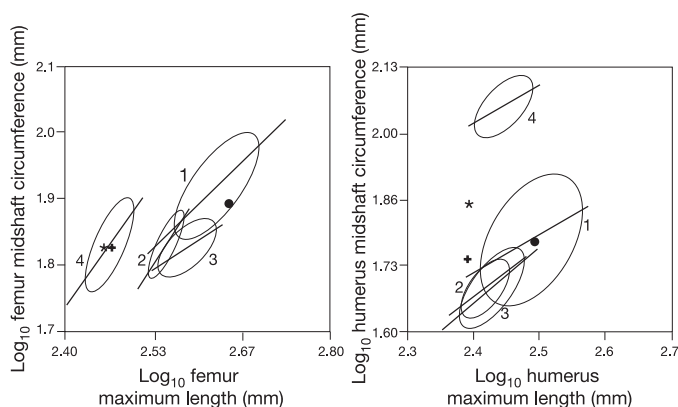


Figure 7 | Femur and humerus shaft robusticity in *H. floresiensis*.

Relationship between mid-shaft circumference and maximum length for the LB1 *H. floresiensis* femur and humerus (star) compared with those for KNM WT-15000 *H. erectus* (filled circle)⁴, AL 288-1 *A. afarensis* (cross)²⁰, modern humans (numbers 1–3) and *Pan paniscus* (number 4). The modern human comparative series is from the robust global sample used previously¹, with the addition of African Pygmies (2) and Andaman Islanders (3). Least squares regression lines and 95% confidence ellipses are shown.

for LB8. However, the tibia is considerably smaller in all dimensions than the holotype and, given the population-specific correlations between limb segment length and stature¹⁶, LB8 was probably a shorter individual than LB1, who had an estimated stature of 106 cm on the basis of femur length¹, a more reliable proxy for stature.

The right humerus, right ulna and left fibula from Spit 58A occurred at the same level as, and immediately adjacent to, the LB1 remains found in Sector VII, and in the same layer of clayey silt. The humerus articulates with the ulna, and the fibula pairs with the right fibula of LB1 excavated in 2003—both left and right fibulas are 226 mm in length. On the basis of size, representation and disposition, all these elements are part of LB1.

The LB1 right humerus is complete, apart from post-mortem damage to the anterior surface of the head and greater tubercle, and has a maximum length of 243 mm. Although the mid-shaft was broken post-mortem, there is no distortion. When compared with a modern human humerus scaled to the same length, the most obvious differences are the greater thickness of the LB1 shaft and the limited degree of humeral torsion (that is, rotation of the humerus head medially relative to the mediolateral axis of the distal end; Fig. 4). In LB1, humeral torsion is approximately 110°, which is the norm for *Hylobates* and quadrupedal primates such as *Macaca*, but is significantly less than in large-bodied apes, modern humans (141–178°) and other known hominins, including *Australopithecus*^{18–20}.

The LB1 right ulna is not as completely preserved as the humerus, and the distal end is not present (Fig. 5). There are relatively few well-preserved Pleistocene or Pliocene hominin ulnae, but comparing the LB1 ulna with those of modern humans, KNM WT-15000 *H. erectus*¹⁴ or AL 288-1 *A. afarensis*²⁰, the most obvious differences are the greater robusticity of the shaft of LB1 and the broader subtrochlear region, with considerably greater areas of attachment for the flexor, pronator and brachialis muscles. The length of the preserved shaft is 190 mm, and comparison with modern human ulnae suggests that, when complete, the ulna would have had a maximum length of 205 mm (see Supplementary Information). The arms of LB1 share the distinctive elongated distal segment common to tropical modern humans^{21,22}, including populations of small average stature, and have an estimated brachial index ((radius length × 100)/humerus length) of 78, which is close to the African human male average²³ (see Supplementary Information).

In association with the previously described femur, tibia and pelvis¹, the humerus and ulna permit an estimate of relative limb and body proportions for the *H. floresiensis* holotype. The only

reported *H. erectus* skeleton, KNM WT-15000, has limb proportions and a body shape similar to *H. sapiens*¹⁴, which differ from the relatively short legs, more teardrop-shaped thoracic cavity and laterally flaring pelvis in AL 288-1 *A. afarensis*^{20,24}. Both the humerus and ulna of LB1 are long relative to femur length (Fig. 6). For example, the humerofemoral index ((humerus length × 100)/femur length) of 85.4 is outside the range of variation for *H. sapiens*, but is the same as AL 288-1 *A. afarensis*, and midway between the indices for apes and humans. The more complete left ilium¹ also indicates that the pelvis is flared antero-laterally, consistent with an australopithecine-shaped thoracic region. Body proportions of LB1 are the same as AL288-1 *A. afarensis*, but differ from all other hominins for which there are reliable data, including *H. erectus*. Abnormal growth seems an unlikely explanation, as growth-hormone-related dwarfism and microcephaly in modern humans result in normal limb and pelvic proportions²⁵.

All the major limb bones of LB1 have shaft and articular surface dimensions that are robust relative to length. For the femur and humerus, midshaft circumference differs from predictions for *H. sapiens* of similar body size, with femur robusticity at the centre of the *Pan paniscus* range of variation, and humerus robusticity midway between the *P. paniscus* and *H. sapiens* ranges (Fig. 7). Thus, estimations of musculature and body mass in *H. floresiensis* would be more accurate if based on chimpanzee rather than human models¹. Although the endocast of LB1 has a gross morphology most similar to *H. erectus*, and distinct from *H. sapiens*, *Australopithecus* and *Pan*²⁶, the encephalization quotient and body-weight-to-brain-weight relationships are similar to *A. afarensis* and *Pan*, not *H. erectus*^{1,27}.

This new evidence has implications for the taxonomy, evolutionary history and physiology of *H. floresiensis*. The cranial morphology, skeletal proportions and endocast of LB1, together with the combination of primitive and derived features evident in both excavated mandibles, confirm that this population of hominins cannot be referred to either *H. sapiens* or *H. erectus*. However, although tooth size and facial morphology dictate inclusion of the species in the genus *Homo*, the genealogy of *H. floresiensis* remains uncertain. Similarities in stature and body proportions with *Australopithecus*, for example, may reflect phylogeny or secondary evolutionary convergence. Either way, *H. floresiensis* is not just an allometrically scaled-down version of *H. erectus*. Other *H. floresiensis* morphological traits, for example in the humeral torsion and ulna, are not shared with any other known hominin species and could have evolved *in situ* as a response to island-specific conditions and associated behavioural changes (see refs 28, 29).

Received 13 January; accepted 12 July 2005.

1. Brown, P. et al. A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**, 1055–1061 (2004).
2. Morwood, M. J. et al. Archaeology and age of a new hominin from Flores in eastern Indonesia. *Nature* **431**, 1087–1091 (2004).
3. Weidenreich, F. The mandible of *Sinanthropus pekinensis*. A comparative study. *Palaeontologica Sinica* **D7**, 1–132 (1936).
4. Kaifu, Y. et al. Taxonomic affinities and evolutionary history of the early Pleistocene hominids of Java: Dentognathic evidence. *Am. J. Phys. Anthropol.* published online 10 March 2005 (doi:10.1002/ajpa.10425).
5. Vekua, A. K. et al. A new skull of early *Homo* from Dmanisi, Georgia. *Science* **297**, 85–89 (2002).
6. Wood, B. A. & Yttererschaut, H. Analysis of the dental morphology of the Plio-Pleistocene hominids. III. Mandibular premolar crowns. *J. Anat.* **154**, 121–156 (1987).
7. Wood, B. A., Abbott, S. A. & Yttererschaut, H. Analysis of the dental morphology of Plio-Pleistocene hominids. IV. Mandibular postcanine root morphology. *J. Anat.* **156**, 107–139 (1988).
8. Gabounia, L., de Lumley, M.-A., Vekua, A., Lordkipanidze, D. & de Lumley, H. Découverte d'un nouvel hominidé à Dmanissi (Transcaucasie, Géorgie). *C.R. Palevol* **1**, 243–253 (2002).
9. Wolpoff, M. H. in *Size and Scaling in Primate Biology* (ed. Jungers, W. L.) 273–318 (Plenum, New York, 1985).
10. Schwartz, J. & Tattersall, I. The human chin revisited: what is it and who has it? *J. Hum. Evol.* **38**, 367–409 (2000).

11. Lavelle, C. L. Investigation of mandibular and neurocranial form. *Acta Anat.* **125**, 238–244 (1986).
12. Lavelle, C. L. A preliminary craniofacial profile evaluation of normo-, micro- and macro-cephalics. *Br. J. Orthod.* **13**, 13–21 (1986).
13. White, T. D. & Johanson, D. C. Pliocene hominid mandibles from the Hadar formation, Ethiopia: 1974–1977 collections. *Am. J. Phys. Anthropol.* **57**, 501–544 (1982).
14. Walker, A. C. & Leakey, R. (eds) *The Nariokotome Homo erectus Skeleton* (Harvard Univ. Press, Cambridge, 1993).
15. Haeusler, M. & McHenry, H. M. Body proportions of *Homo habilis* reviewed. *J. Hum. Evol.* **46**, 433–465 (2004).
16. Oliver, G. The stature of *Australopithecus*. *J. Hum. Evol.* **5**, 529–534 (1976).
17. Hens, S. M., Konigsberg, L. W. & Jungers, W. L. Estimating stature in fossil hominids: which regression model and reference sample to use. *J. Hum. Evol.* **38**, 767–784 (2000).
18. Evans, F. G. & Krah, V. E. The torsion of the humerus: a phylogenetic study from fish to man. *Am. J. Anat.* **76**, 303–337 (1945).
19. Larson, S. G. Subscapularis function in gibbons and chimpanzees: implications for interpretation of humeral head torsion in hominoids. *Am. J. Phys. Anthropol.* **76**, 449–462 (1988).
20. Johanson, D. C. *et al.* Morphology of the Pliocene partial hominid skeleton (A.L. 288–1) from the Hadar Formation, Ethiopia. *Am. J. Phys. Anthropol.* **57**, 403–451 (1982).
21. Roberts, D. F. in *Addison-Wesley Module in Anthropology* 1–38 (Addison-Wesley, Reading, 1973).
22. Ruff, C. B. Morphological adaptation to climate in modern and fossil hominids. *Yb. Phys. Anthropol.* **37**, 65–107 (1994).
23. Aiello, A. & Dean, C. *An Introduction to Human Evolutionary Anatomy* (Academic, London, 1990).
24. Schmid, P. Eine rekonstruktion des skelettes von A.L. 288–1 (Hadar) und deren konsequenzen. *Folia Primatol. (Basel)* **40**, 283–306 (1983).
25. Hiratal, G. M., Silverman, B. L., Dupuis, J. & Baumann, G. Phenotype and genetic analysis of a syndrome caused by an inactivating mutation in the growth hormone-releasing hormone receptor: Dwarfism of Sindh. *J. Clin. Endocrinol. Metab.* **83**, 4065–4074 (1998).
26. Falk, D. *et al.* The brain of LB1, *Homo floresiensis*. *Science* **308**, 242–245 (2005).
27. Martin, R. D. Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature* **293**, 57–60 (1981).
28. Köhler, M. & Moyà-Solà, S. Reduction of brain and sense organs in the fossil insular bovid *Myotragus*. *Brain Behav. Evol.* **63**, 125–140 (2004).

29. van der Geer, A. A. E. On the astragalus of the Miocene endemic deer *Hoplitomeryx* from the Gargano (Italy). *Deinsea* **7**, 325–336 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The 2004 excavations at Liang Bua were funded by a Discovery Project grant from the Australian Research Council (ARC) to M.J.M. The ARC also supported R.G.R. through a Senior Research Fellowship. K.E.W. was in receipt of postgraduate scholarships from the University of Wollongong, and P.B. and T.M. received a grant from the University of New England. Excavations were undertaken under Asisten Deputi Urusan Arkeologi Nasional, Number 1100/SB/U.ARNAS/VII/04, and Pemerintah Kabupaten Manggarai, Surat Rekomendasi No: 565/Kesbang.IV/VII/2004. R. P. Soejono was the Indonesian Institutional Counterpart. Other participants included C. Lentfer, G. van den Bergh, C. Turney, D. Hobbs, K. Grant, D. Y. Yuniawati, A. Brumm, Rikus, Deus, Leo, Ansel, Agus, Seus, Camellus, Gaba, Rius, Beni and Piet. P. Jordan, D. Falk, S. G. Larson and K. Morwood commented on an earlier version of the paper, and D. Hobbs redrafted Fig. 1. P.B. and T.M. thank R. Kruszynski (Natural History Museum, London), R. Orban (Institut Royal des Sciences Naturelles de Belgique, Bruxelles) and W. Wendelen (Royal Museum of Central Africa, Tervuren) for providing access to skeletal collections in their care.

Author Contributions M.J.M. was Chief Investigator and the Australian Institutional Counterpart in this ARC project. P.B. was responsible for analysis and interpretation of the hominin remains. T.S., J., E.W.S., R.A.D. and S.W. directed aspects of the excavations, recording and analyses. Thermoluminescence ages were provided by R.G.R. and K.E.W., who also described the stratigraphy. T.M. and P.B. collected comparative data on small-bodied modern humans and non-human primates. As Director of the Indonesian Centre for Archaeology, T.D. authorized the excavations, provided support and assisted with stratigraphic interpretation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.J.M. (mmorwood@pobox.une.edu.au) and P.B. (pbrown3@pobox.une.edu.au).

LETTERS

Vertebrate Smoothened functions at the primary cilium

Kevin C. Corbit^{1,2}, Pia Aanstad^{1,2,3}, Veena Singla^{1,2}, Andrew R. Norman^{1,2}, Didier Y. R. Stainier^{1,2,3}
& Jeremy F. Reiter^{1,2}

The unanticipated involvement of several intraflagellar transport proteins in the mammalian Hedgehog (Hh) pathway has hinted at a functional connection between cilia and Hh signal transduction^{1,2}. Here we show that mammalian Smoothened (Smo), a seven-transmembrane protein essential for Hh signalling³, is expressed on the primary cilium. This ciliary expression is regulated by Hh pathway activity; Sonic hedgehog or activating mutations in Smo promote ciliary localization, whereas the Smo antagonist cyclopamine inhibits ciliary localization. The translocation of Smo to primary cilia depends upon a conserved hydrophobic and basic residue sequence homologous to a domain previously shown to be required for the ciliary localization of seven-transmembrane proteins in *Caenorhabditis elegans*⁴. Mutation of this domain not only prevents ciliary localization but also eliminates Smo activity both in cultured cells and in zebrafish embryos. Thus, Hh-dependent translocation to cilia is essential for Smo activity, suggesting that Smo acts at the primary cilium.

Nearly all interphase vertebrate cells possess a primary cilium that extends into the extracellular environment⁵. Recent studies have pointed to a role for primary cilia as sensors capable of transmitting diverse types of information from the environment⁶. One example of this is the photoreceptor outer segment, a derivative of the primary cilium and the site at which light activates the seven-transmembrane (7TM) protein rhodopsin⁷. This connection between 7TM proteins and cilia seems to be an ancient one, as odorant receptors function on the cilia of *C. elegans* sensory neurons⁴.

Mutations in genes required for cilia formation cause defects in mouse neural tube patterning indicative of altered Hh signalling^{1,2}. The mechanism by which these genes participate in Hh signalling is unclear, but one possibility is that their involvement signifies a requirement for cilia in Hh signal transduction.

Hh signalling in *Drosophila*, zebrafish and mice is transduced through Smo, a 7TM protein³. To investigate the expression of endogenous Smo during mouse development we made use of two polyclonal anti-Smo antibodies. The antibodies are highly specific, as no protein is detected in *Smo*^{-/-} samples by immunofluorescence and western blot analyses (Supplementary Fig. 1). During early somite stages, Smo protein is modestly upregulated in the cells of the node (Fig. 1a), an organizer involved in several developmental processes including left-right axis determination and floor plate induction.

One structural characteristic that distinguishes ventral node cells from other cells of the early embryo is the presence of a primary cilium⁸. Examination of Smo localization in the nodes of early headfold, late headfold, one-somite, three-somite and five-somite-stage embryos revealed that Smo is located predominantly on

primary cilia (Fig. 1b, c). Although Smo does not localize uniformly along the entire length of cilia (Fig. 1d), average expression on cilia is greater than threefold higher than elsewhere on nodal cells.

The increased level of Smo protein expression specifically in the node is notable given that *Smo* messenger RNA is present throughout

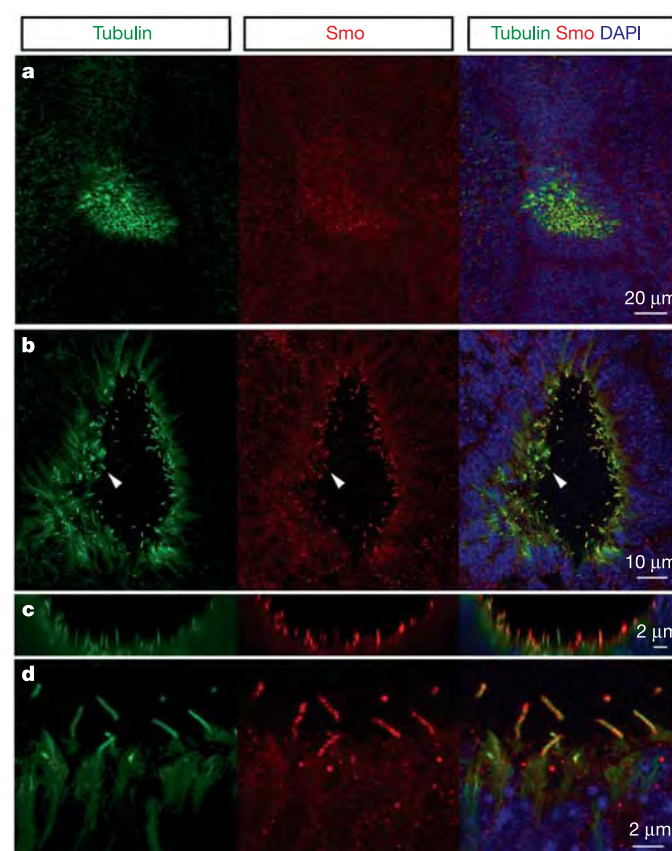


Figure 1 | Smo is expressed on nodal primary cilia. Expression of the ciliary marker acetylated Tubulin (green) and Smo (red). Nuclei are visualized with DAPI staining (blue). **a**, Ventral view of the node of a three-somite embryo with anterior oriented upwards. Elevated levels of acetylated Tubulin staining demarcate the ventral node cells. Smo expression in the node is approximately 170% that of surrounding cells. **b**, A higher magnification ventral view of a node demonstrating ciliary Smo co-localization with acetylated Tubulin (arrowhead). **c**, Optical cross-section through the node. **d**, High-magnification view of nodal primary cilia.

¹Developmental and Stem Cell Biology Program, ²Diabetes Center, and ³Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA.

the embryo⁹. Because *Sonic hedgehog* and *Indian hedgehog* genes (*Shh* and *Ihh*) are expressed in and near the node^{9,10}, respectively, we hypothesized that, as in *Drosophila*^{11,12}, mouse Smo might be stabilized in regions of Hh signalling. Indeed, Smo is upregulated specifically in and around the Shh-producing cells of the node, as demonstrated by the co-expression of Smo and Shh in early headfold-stage embryos (Supplementary Fig. 2).

To test whether Smo localization to cilia is regulated by Hh signalling, we created a renal epithelial MDCK (Madin–Darby canine kidney) cell line constitutively expressing Myc-tagged murine Smo. Smo is detectable on the plasma cell membrane and intracellular vesicles of MDCK cells (Fig. 2a), as previously described for *Drosophila* salivary gland and imaginal disc cells and rat KNRK (Kirsten

murine sarcoma virus-transformed normal rat kidney) cells^{11–14}. We examined Smo distribution in MDCK cells cultured in the presence of Shh or cyclopamine, a Smo antagonist¹⁵. After culture for 1 h in Shh-conditioned medium, the amount of Smo present on the primary cilium is markedly upregulated (Fig. 2b). Similar localization of endogenous Smo is seen in Shh-treated mouse embryonic fibroblasts (MEFs) and IMCD-3 (inner medullary collecting duct) kidney cells (Supplementary Fig. 4). Addition of cyclopamine to Shh-conditioned medium eliminates detectable Smo from the primary cilium of MDCK cells (Fig. 2c).

The mouse homologue of an activated Smo mutant, SmoA1, has previously been demonstrated to affect Smo distribution¹⁴. Unlike wild-type Smo and unlike its behaviour in unciliated cells¹⁵, SmoA1 strongly localizes to the primary cilium of MDCK cells even in the absence of Shh (Fig. 2c). High concentrations of cyclopamine sufficient to prevent SmoA1 activation of the Hh pathway¹⁵ are able to eliminate SmoA1 from the cilium (Fig. 2d). To assess whether the presence of Smo on nodal cilia is also regulated by Hedgehog signalling, we cultured mouse embryos in cyclopamine. Comparison of the treated and untreated embryos reveals that cyclopamine markedly diminishes Smo localization to cilia (Fig. 2g, h), demonstrating that inhibition of Hh signalling *in vivo* correlates with the absence of Smo from the cilium.

The upregulation of Smo on the primary cilium in response to Shh or activating mutations suggests that the regulation of mammalian Smo trafficking may be more similar to that of *Drosophila* Smo than previously recognized. In both a *Drosophila* cell line and in imaginal discs, Hh induces the accumulation of Smo at the cell surface^{11,12}, whereas in rat KNRK cells, Shh stimulation induces the internalization of cell-surface Smo¹⁴. One reason KNRK cells process Smo differently may be that they lack cilia (J.F.R., data not shown). Our data suggest that Hh-dependent cell-surface enrichment of Smo may be common to both *Drosophila* cells and ciliated mammalian cells.

Ciliary localization of the *C. elegans* 7TM proteins ODR-10 and STR-1 depends on a hydrophobic and basic residue motif immediately carboxy-terminal to the seventh transmembrane segment¹⁶. This ciliary localization motif is also present in other 7TM proteins known to localize to mammalian primary cilia, such as somatostatin receptor 3 and serotonin receptor 6 (refs 17, 18). Sequence inspection reveals that Smo proteins also contain a hydrophobic and basic

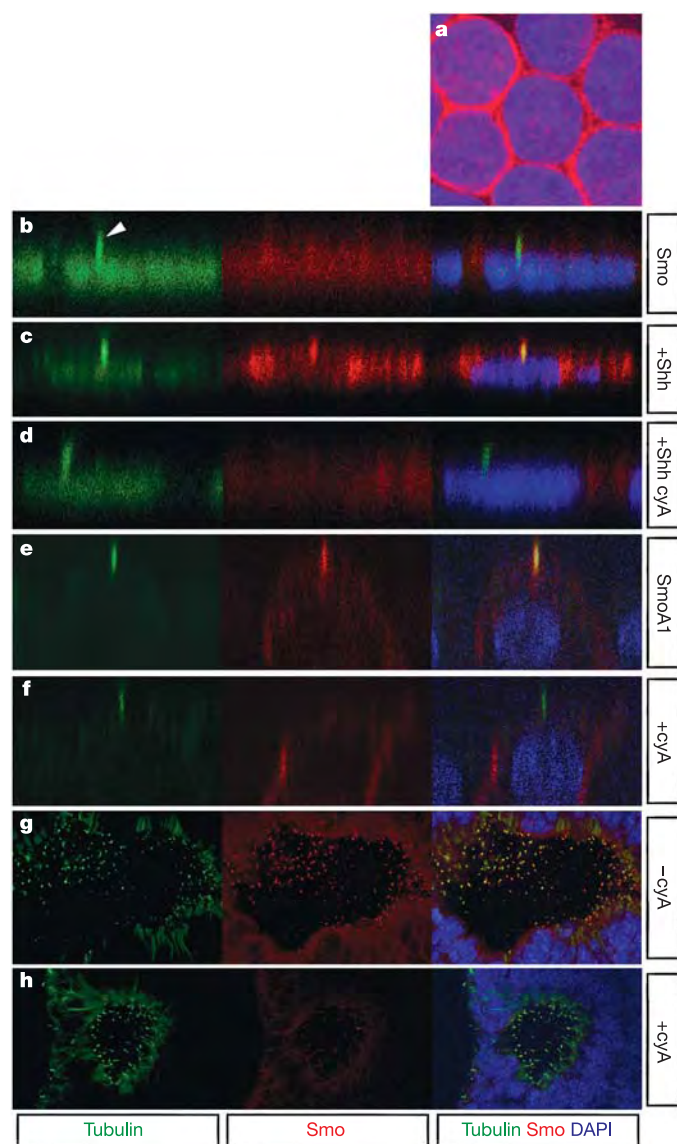


Figure 2 | Hh pathway activity regulates the ciliary localization of Smo. **a–f**, MDCK cell expression of acetylated tubulin (green) and the Myc tag of Smo (red). Nuclei are visualized with DAPI (blue). **a**, Transverse optical section of cells expressing wild-type Smo. **b–f**, Apical–basal optical sections. **b**, Cell expressing wild-type Smo. Acetylated tubulin expression marks the primary cilium (arrowhead). **c**, Smo-expressing cell cultured in the presence of Shh. **d**, Smo-expressing cell treated with both Shh and cyclopamine (cyA). **e**, Cell expressing SmoA1. **f**, SmoA1-expressing cell cultured in the presence of cyclopamine. **g, h**, Expression of acetylated tubulin (green) and Smo (red) in the nodes of headfold-stage embryos cultured in the absence (**g**) or presence (**h**) of cyclopamine. Nuclei are visualized with DAPI (blue).

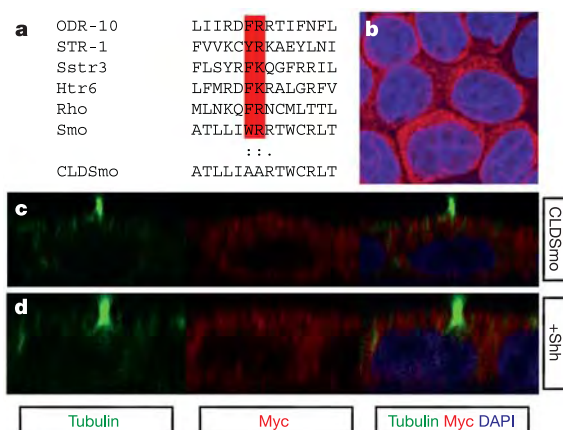


Figure 3 | Identification of a conserved Smo ciliary localization motif. **a**, Sequence comparison of the carboxy-terminal juxtamembrane sequences of 7TM proteins known to localize to cilia. ODR-10 and STR-1 are *C. elegans* 7TM receptors. Murine sequences are included for somatostatin receptor 3 (Sstr3), serotonin receptor 6 (Htr6), rhodopsin (Rho) and Smo. The conserved hydrophobic and basic residue motif is highlighted in red. CLDSmo is a mutant Smo lacking this motif. **b–d**, CLDSmo-expressing MDCK cells stained as in Fig. 2. **b**, Transverse optical section. Compare to Fig. 2a. **c**, Apical–basal optical section. **d**, Apical–basal optical section of a cell cultured in the presence of Shh.

residue motif carboxy-terminal to the seventh transmembrane segment (Fig. 3a).

To test whether this motif is required for ciliary localization, we replaced Trp 549 and Arg 550 of mouse Smo-Myc with alanines (Fig. 3a). The resultant mutant version of Smo is expressed on the cell plasma membrane and intracellular vesicles of MDCK cells, suggesting that it is folded and at least partly released from the endoplasmic reticulum (Fig. 3b). However, unlike wild-type Smo, this mutant Smo is absent from primary cilia, even in the presence of Shh (Fig. 3c, d). Consequently, we have named this mutant CLDSmo for ciliary localization defective.

These results indicate that the conserved motif carboxy-terminal to the last transmembrane segment of mouse Smo is a motif essential for ciliary localization, providing evidence that, like *C. elegans* 7TM proteins, mammalian 7TM proteins can utilize a juxtamembrane hydrophobic and basic residue sequence for proper localization to cilia. *Drosophila* Smo contains a similar motif, although cilia do not seem to be required for *Drosophila* Hh signal transduction^{19,20}. It is tempting to speculate that this domain may participate in the movement of *Drosophila* Smo to the plasma membrane in response to Hh signalling^{11,12}.

To test whether localization to primary cilia is required for Smo function, we compared the activities of wild-type Smo and CLDSmo in two established read-outs of Hh pathway activation. First, we determined Hh responsiveness in the presence of wild-type or CLDSmo using a Gli-dependent luciferase reporter in NIH-3T3 cells²¹. In the absence of transfected Smo, Shh-conditioned medium stimulates a threefold induction of normalized luciferase activity (Fig. 4a). Cells transfected with wild-type Smo exhibit a robust increase in luciferase activity, as previously reported²¹. In contrast,

CLDSmo does not stimulate Hh pathway activity above control levels.

Second, we compared the function of wild-type Smo and CLDSmo in zebrafish. Hh signalling is required for specification of the zebrafish slow muscle lineage^{22–24}. Within the slow muscle lineage, Prox1-expressing superficial slow fibres (SSFs) require low to medium levels of Hh pathway activity, whereas Engrailed-expressing muscle pioneers require high levels of Hh pathway activity²⁵. Disruption of the zebrafish Smo homologue, *slow muscle omitted* (*smu*), prevents formation of both Prox1-positive SSFs and Engrailed-positive muscle pioneers, resulting in embryos that exhibit characteristic U-shaped somites (Fig. 4b). Injection of wild-type murine Smo mRNA into *smu* mutant embryos rescued this morphological defect and restored formation of both SSFs and muscle pioneers (Fig. 4b–d; see also Supplementary Table). In contrast, injection of CLDSmo mRNA into *smu* mutant embryos did not rescue shape, SSF formation or muscle pioneer formation (Fig. 4b–d; see also Supplementary Table). These results demonstrate that the ciliary localization motif is required for Smo function. Whether this motif might have roles in addition to ciliary localization remains to be determined.

Thus, not only does Smo ciliary localization depend upon Hh signalling, but Hh signalling depends upon a Smo ciliary localization motif. Taken together, these findings strongly suggest that Smo acts at the primary cilium to transduce Hh signalling and that Smo localization to cilia is a key regulated step in Hh pathway activation. Our studies provide an example of a 7TM protein translocating to the primary cilium in response to pathway activation, and it will be interesting to determine whether the localization of other ciliary 7TM proteins is similarly regulated.

Previous studies have suggested that redistribution of Smo to the appropriate subcellular location might be the main determinant of Smo activity¹³. Our findings indicate that in mammals, the primary cilium may be that subcellular location. However, whether Smo functions at the cilium in all cell types remains to be determined. In addition, how Smo activates the Hh pathway at the cilium remains unclear. One possibility is that Smo must translocate to the cilium to encounter an activator. An alternative possibility is that the cilium structurally coordinates Smo with downstream Hh pathway components to permit productive interactions. In support of this second possibility, β -arrestin 2, a component of the zebrafish Hh signal transduction apparatus and a known Smo interactor^{26,27}, is concentrated in cilia²⁸.

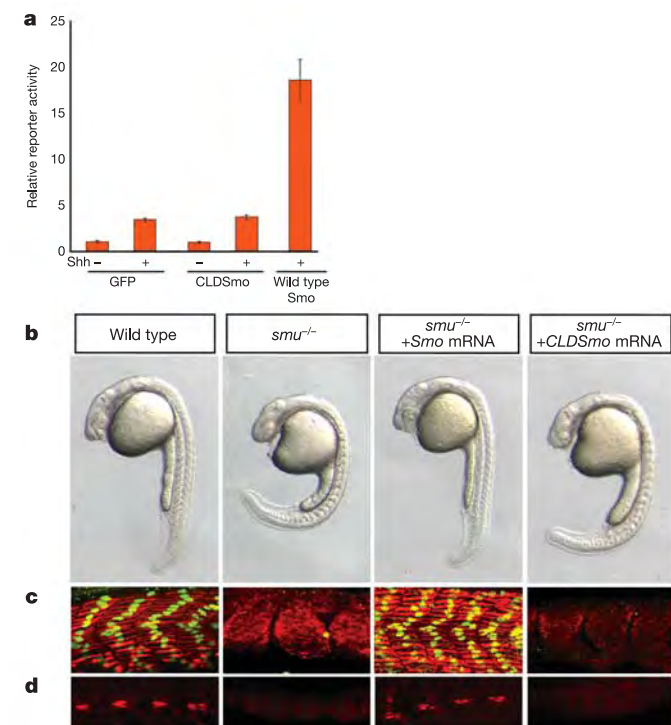


Figure 4 | CLDSmo is inactive. **a**, Induction of a Hh-responsive Gli-luciferase reporter in NIH-3T3 cells. Unlike wild-type Smo, CLDSmo expression fails to either activate the pathway or alter Hh responsiveness. Error bars indicate 1 s.d. **b**, Lateral views of 24 h post-fertilization zebrafish embryos. Shown are a wild-type control embryo and three *smu* mutants, two of which have been injected with 250 pg mRNA encoding either wild-type Smo or CLDSmo. **c**, F59 staining of slow muscle fibres (red), and Prox1-expressing superficial slow fibres (green). **d**, Engrailed (Eng) staining of muscle pioneers.

METHODS

Immunohistochemistry. Mouse embryos and MDCK cells were fixed with 4% PFA for 1 h at 4 °C, washed in PBT (PBS plus 0.2% Triton-X100), and blocked with PBT, 2% BSA and 1% goat serum for 1 h at 4 °C. Primary antibodies were added in block buffer and incubated overnight at 4 °C. Primary antibodies used in this study are anti-Myc (Abcam ab9106, 1:500), mouse anti-acetylated Tubulin (Sigma 6-11B-1, 1:1,000) and rabbit anti-Smo (Lifespan Biosciences A2666 and A2668, 1:200). The anti-Smo antibodies perform equivalently. Subsequently, cells were washed in PBT and incubated in a secondary block (PBT plus 10% donkey serum and 2% BSA) at room temperature for 1 h, followed by incubation with secondary antibodies (1:400 donkey anti-rabbit IgG-Alexa 594 and donkey anti-mouse IgG-Alexa 488; Molecular Probes) in PBT plus 2% BSA at room temperature for 1 h. After washes with PBT and 10 min 4',6-diamidino-2-phenylindole (DAPI) incubation, the samples were mounted and imaged with a Leica DMIRE2 confocal microscope. Images were processed using Leica Confocal, Metamorph 6.1 and Velocity 3.1 software. Zebrafish embryos were fixed overnight at 4 °C in 3% PFA, and staining was performed in PBT-BSA (PBS, 0.3% Triton X-100, 4% BSA, 0.02% Na₂S₂O₈) using F59 anti-Myosin heavy chain (gift from F. Stockdale), anti-Engrailed (gift from S. Amacher) and anti-Prox1 antibodies (Chemicon International). Zebrafish embryos were visualized with a Zeiss LSM5 Pascal confocal microscope.

Expression constructs. Generation of the CLDSmo expression construct, pGE-CLDSmo-Myc-His, was engineered with the QuikChange II XL site-directed mutagenesis kit (Stratagene) using pGE-Smo-Myc-His (gift from P. Beachy) as template. The sequence of the pGE-CLDSmo-Myc-His open reading frame was confirmed.

Cell culture. MDCK cells were a gift from K. Mostov and were grown in MEM, 5% FBS, penicillin and streptomycin. To create MDCK lines stably expressing wild-type Smo or CLDSmo, the cells were transfected using Lipofectamine2000 (Invitrogen) and split 1:10 the following day. Beginning 36 h after transfection, cells were selected in 600 µg ml⁻¹ neomycin, and 5 days later, individual clones were isolated and expanded. The expression of wild-type Smo and CLDSmo was confirmed by western blot with goat anti-Myc antibody (Novus Biologicals). MDCK cell lines expressing comparable levels of Smo or CLDSmo were seeded at 3 × 10⁵ cells per transwell filter (12 mm filter diameter/0.4 µm pore size, Costar) and fed daily for 6 days. Experimental samples were incubated in 10 µM cyclopamine, Shh-conditioned media, or both. Shh-conditioned media was generated as previously described²⁹.

Murine embryo culture. Embryonic day (E)8.0 embryos were isolated in HEPES-buffered DMEM plus 10% FBS, and then cultured in DR75 medium in the presence or absence of 10 µM cyclopamine. Embryos were cultured statically for 1 h at 37 °C in an incubator gassed with 5% CO₂ in air before being processed for immunohistochemistry as above.

Gli-luciferase reporter assay. NIH-3T3 cells were plated at 10⁵ cells per well in 24-well plates. On the next day, we transfected cells with *Renilla* luciferase (pRL-TK, Promega), Gli-luciferase reporter, and a Smo expression construct using Eugene6 (Roche) as previously described²¹. Two days later, cells were moved to low serum media (DMEM, 2% delipidated fetal calf serum, ITS supplement, 5 mM HEPES, penicillin and streptomycin) for 8 h, then treated with Shh-conditioned media for 18 h. All reporter assays were normalized for transfection efficiency using *Renilla* luciferase control (Dual luciferase assay, Promega).

Zebrafish mRNA injection. Embryos from *smu*^{b641} heterozygote intercrosses were injected at the one- or two-cell stage with 250 pg mRNA encoding either mouse wild-type Smo or CLDSmo, and raised at 28 °C until 24 h after fertilization.

Genotyping. Mouse *Smo* embryos were genotyped as described⁹. Zebrafish *smu*^{b641} embryos were genotyped by polymerase chain reaction (PCR) amplification (primers 5'-TTACGTCAACGCCTGTTCTTCCTT-3' and 5'-TCGCA-TAGTGTGTCGCTTT-3'), followed by *DdeI* restriction enzyme digestion.

Received 26 April; accepted 9 August 2005.

Published online 31 August 2005.

- Huangfu, D. et al. Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* **426**, 83–87 (2003).
- Liu, A., Wang, B. & Niswander, L. A. Mouse intraflagellar transport proteins regulate both the activator and repressor functions of Gli transcription factors. *Development* **132**, 3103–3111 (2005).
- Hooper, J. E. & Scott, M. P. Communicating with Hedgehogs. *Nature Rev. Mol. Cell Biol.* **6**, 306–317 (2005).
- Dwyer, N. D., Troemel, E. R., Sengupta, P. & Bargmann, C. I. Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell* **93**, 455–466 (1998).
- Wheatley, D. N. Primary cilia in normal and pathological tissues. *Pathobiology* **63**, 222–238 (1995).
- Salisbury, J. L. Primary cilia: putting sensors together. *Curr. Biol.* **14**, R765–R767 (2004).
- Takemoto, D. J. & Cunnick, J. M. Visual transduction in rod outer segments. *Cell. Signal.* **2**, 99–104 (1990).
- Sulik, K. et al. Morphogenesis of the murine node and notochordal plate. *Dev. Dyn.* **201**, 260–278 (1994).
- Zhang, X. M., Ramalho-Santos, M. & McMahon, A. P. Smoothed mutants reveal redundant roles for Shh and Ihh signalling including regulation of L/R symmetry by the mouse node. *Cell* **106**, 781–792 (2001).
- Roelink, H. et al. Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog expressed by the notochord. *Cell* **76**, 761–775 (1994).
- Zhu, A. J., Zheng, L., Suyama, K. & Scott, M. P. Altered localization of *Drosophila* Smoothed protein activates Hedgehog signal transduction. *Genes Dev.* **17**, 1240–1252 (2003).
- Denef, N., Neubuser, D., Perez, L. & Cohen, S. M. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothed. *Cell* **102**, 521–531 (2000).
- Nakano, Y. et al. Functional domains and sub-cellular distribution of the Hedgehog transducing protein Smoothed in *Drosophila*. *Mech. Dev.* **121**, 507–518 (2004).
- Incardona, J. P., Gruenberg, J. & Roelink, H. Sonic hedgehog induces the segregation of patched and smoothed in endosomes. *Curr. Biol.* **12**, 983–995 (2002).
- Chen, J. K., Taipale, J., Cooper, M. K. & Beachy, P. A. Inhibition of Hedgehog signalling by direct binding of cyclopamine to Smoothed. *Genes Dev.* **16**, 2743–2748 (2002).
- Dwyer, N. D., Adler, C. E., Crump, J. G., L'Etoile, N. D. & Bargmann, C. I. Polarized dendritic transport and the AP-1 mu1 clathrin adaptor UNC-101 localize odorant receptors to olfactory cilia. *Neuron* **31**, 277–287 (2001).
- Handel, M. et al. Selective targeting of somatostatin receptor 3 to neuronal cilia. *Neuroscience* **89**, 909–926 (1999).
- Brailov, I. et al. Localization of 5-HT(6) receptors at the plasma membrane of neuronal cilia in the rat brain. *Brain Res.* **872**, 271–275 (2000).
- Ray, K. et al. Kinesin-II is required for axonal transport of choline acetyltransferase in *Drosophila*. *J. Cell Biol.* **147**, 507–518 (1999).
- Han, Y. G., Kwok, B. H. & Kernan, M. J. Intraflagellar transport is required in *Drosophila* to differentiate sensory cilia but not sperm. *Curr. Biol.* **13**, 1679–1686 (2003).
- Taipale, J. et al. Effects of oncogenic mutations in Smoothed and Patched can be reversed by cyclopamine. *Nature* **406**, 1005–1009 (2000).
- Varga, Z. M. et al. Zebrafish smoothed functions in ventral neural tube specification and axon tract formation. *Development* **128**, 3497–3509 (2001).
- Chen, W., Burgess, S. & Hopkins, N. Analysis of the zebrafish smoothed mutant reveals conserved and divergent functions of hedgehog activity. *Development* **128**, 2385–2396 (2001).
- Barresi, M. J., Stickney, H. L. & Devoto, S. H. The zebrafish slow-muscle-omitted gene product is required for Hedgehog signal transduction and the development of slow muscle identity. *Development* **127**, 2189–2199 (2000).
- Wolff, C. et al. *iguana* encodes a novel zinc-finger protein with coiled-coil domains essential for Hedgehog signal transduction in the zebrafish embryo. *Genes Dev.* **18**, 1565–1576 (2004).
- Wilbanks, A. M. et al. β-arrestin 2 regulates zebrafish development through the hedgehog signalling pathway. *Science* **306**, 2264–2267 (2004).
- Chen, W. et al. Activity-dependent internalization of smoothed mediated by β-arrestin 2 and GRK2. *Science* **306**, 2257–2260 (2004).
- Dawson, T. M. et al. β-adrenergic receptor kinase-2 and β-arrestin-2 as mediators of odorant-induced desensitization. *Science* **259**, 825–829 (1993).
- Cooper, M. K., Porter, J. A., Young, K. E. & Beachy, P. A. Teratogen-mediated inhibition of target tissue response to Shh signalling. *Science* **280**, 1603–1607 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to F. Schaufele for microscopy assistance, and K. Liu for advice about MDCK cell culture. This work was assisted by valuable conversations with P.-T. Chuang, W. Marshall and A. Peterson. Support for microscopy was provided by the NIH. Other support was provided by the NIH (D.Y.R.S.), the Sandler Foundation and a March of Dimes Basil O'Connor Award (J.F.R.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.F.R. (jreiter@diabetes.ucsf.edu).

LETTERS

SERRATE coordinates shoot meristem function and leaf axial patterning in *Arabidopsis*

Stephen P. Grigg¹, Claudia Canales¹, Angela Hay¹ & Miltos Tsiantis¹

Leaves of flowering plants are determinate organs produced by pluripotent structures termed shoot apical meristems. Once specified, leaves differentiate an adaxial (upper) side specialized for light capture, and an abaxial (lower) side specialized for gas exchange. A functional relationship between meristem activity and the differentiation of adaxial leaf fate has been recognized for over fifty years, but the molecular basis of this interaction is unclear. In *Arabidopsis thaliana*, activity of the class I KNOX (*KNOTTED1*-like homeobox) genes *SHOOTMERISTEMLESS* (*STM*) and *BREVIPEDICELLUS* (*BP*) is required for meristem function but excluded from leaves^{1–3}, whereas members of the HD-Zip III (class III homeodomain leucine zipper) protein family function to promote both meristem activity and adaxial leaf fate^{4–6}. Here we show that the zinc-finger protein SERRATE acts in a microRNA (miRNA) gene-silencing pathway to regulate expression of the HD-Zip III gene *PHABULOSA* (*PHB*) while also limiting the competence of shoot tissue to respond to KNOX expression. Thus, SERRATE acts to coordinately regulate meristem activity and leaf axial patterning.

Coordination of meristem activity with the differentiation of adaxial cell types was first suggested by observations that surgical separation of the shoot apical meristem (SAM) from developing leaf primordia prevents elaboration of adaxial identity in leaves⁷. Insights into the molecular basis of this interaction have been provided by two observations. First, when KNOX proteins that are required for meristem development^{1–3} are expressed in the leaf, ectopic meristems form exclusively on the adaxial leaf surface^{8,9}. Second, HD-Zip III gene family members such as *PHB* are required for meristem development, but their expression persists in developing leaf primordia, where they promote adaxial cell fate^{4,5}. Broadened and elevated expression of *PHB* in *phb-d* mutants results in increased SAM size, ectopic meristem formation and leaf adaxialization^{4,10,11} associated with hypomethylation of the *PHB* locus^{4,10,11}. Thus, repression of *PHB* expression is critical for coordinately regulating development of the SAM and adaxial leaf domain. Identifying the factors implementing this repression as well as those responsible for the integration of KNOX and *PHB* activities is therefore crucial for elucidating the processes that coordinate development of the SAM and adaxial leaf domain.

To understand the interplay between meristem function and differentiation of adaxial cell types, we isolated new alleles at the *SERRATE* (*SE*) locus. *SE* is a good candidate for regulating this process because it is expressed in both the SAM and adaxial leaf domain¹². Additionally, the weak *se-1* mutant allele pleiotropically perturbs both shoot and leaf development^{12,13}. Although the precise mode of *SE* action is unclear, sequence analysis indicates limited similarity to zinc-finger proteins that regulate chromatin activity¹². Genes distantly related to *SE* also exist in animal genomes, where they may have crucial developmental functions, as suggested by an embryo-lethal phenotype conditioned by mutation of a *SE*-like

gene in zebrafish^{14,15}. Characterization of two new *se* alleles *se-2* and *se-3* (Supplementary Fig. S1) revealed phenotypes not seen in *se-1* mutants, including adaxial leaf curling, loss of asymmetric differentiation of abaxial and adaxial cell types, and development of trumpet-shaped or radial leaves (Fig. 1a–g). Vascular polarity of *se-3* leaves was also perturbed, with xylem elements forming both adaxially and abaxially in the vascular bundle; this is in contrast to wild-type (Columbia ecotype) plants, in which xylem elements only form adaxially (Fig. 1f, g). In radial *se-3* leaves, vascular bundles either failed to differentiate or consisted solely of xylem elements (Fig. 1h).

Therefore, the *se* phenotype suggests that *SE* acts either to limit the activity of adaxial-promoting factors in the shoot, or to promote the activity of abaxial-promoting factors. All three *se* alleles show increased expression of the two redundantly acting HD-Zip III genes *PHB* and *PHAVOLUTA* (*PHV*) (Fig. 2a, Supplementary Fig. S2 and data not shown). These data suggest that *SE* functions to limit activity of these HD-Zip III genes and hence correctly define the adaxial domain in leaves. This suggestion is consistent with the observation that, similar to *phb-1d* and *phv-1d* mutants, *se-2* and *se-3* plants show enlarged embryonic SAMs (Fig. 1i, j). Additionally, *se-3* mutants develop a substantially hypertrophic vegetative SAM at a frequency of 16% (Fig. 1k, l and data not shown). Thus, *SE* activity is required for both correct axial patterning of leaves and control of SAM size.

To determine whether altered HD-Zip III gene expression patterns are consistent with aspects of the *se* mutant phenotype, we examined the localization of *PHB* transcripts *in situ*. In contrast to wild-type *Arabidopsis*, in which *PHB* is restricted to the adaxial side of the leaf⁴, in *se* plants it is expressed broadly, extending into the abaxial domain (arrowheads in Fig. 2b–d). Furthermore, the *PHB* expression domain is expanded in *se* mutant SAMs (Fig. 2b, c), and *se-3* plants with hypertrophic apices express *PHB* transcript strongly throughout multiple foci at the flanks of the SAM (Fig. 2h). Analysis of *STM* expression indicated that these foci correspond to sites of *STM* downregulation (Fig. 2e–g), and are therefore likely to represent lateral organ primordia. Misexpression of *PHB* in *se* mutants was evident at the eight-cell stage of embryogenesis (Fig. 2i–l). Thus, *SE* defines the expression level and domain of the *PHB* gene during *Arabidopsis* embryonic and vegetative development. However, it is unlikely that *SE* is a conventional transcriptional repressor of *PHB*, because overexpression of HD-Zip III genes fails to perturb leaf or shoot development unless a miRNA complementary site present in the transcript is mutated^{5,16}. Therefore, *SE* may regulate *PHB* in a miRNA-dependent process.

To test whether *SE* is required *in trans* for miRNA-based regulation of *PHB*, we expressed *PHB* in *se-2* plants under the broadly expressed CaMV 35S promoter. Unlike *se-3* plants, *se-2* plants do not show severe axial patterning phenotypes. We therefore predicted that if *SE* regulates *PHB* through a miRNA-dependent pathway, then *PHB*

¹Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK.

overexpression in this background should elicit strong phenotypes similar to dominant *phb-d* mutants^{4,10}. Analysis of transgenic plants confirmed this prediction, as the majority of *35S::PHB; se-2* plants (38/57) showed severe patterning defects, including multiple radial leaves, ectopic meristems (Fig. 3a, b) and growth arrest at the seedling apex (data not shown). Neoplastic growth and fasciation were also observed in these plants (Fig. 3c and data not shown). Consistent with previous reports, *35S::PHB* had little or no effect on wild-type development (data not shown and refs 4, 5, 16). These observations suggest that SE acts in a miRNA-dependent pathway to repress *PHB*.

To test directly the involvement of SE in miRNA-dependent *PHB* regulation, we assayed profiles of transcripts generated from the *PHB* locus in wild-type and *se* mutant plants using the 5' rapid amplification of cDNA ends (RACE) technique. In a wild-type background, we detected *PHB* transcripts cleaved in the miRNA complementary site (Fig. 3d). Such products were not detected in *se-3* tissue, indicating that miRNA-based regulation of *PHB* is sensitive to the loss of SE activity. To understand the nature of this regulation, we examined levels of both the miR165 and miR166 primary transcripts and the mature 21-nucleotide (nt) transcripts that are complementary to *PHB* mRNA. Levels of miR165 and miR166 primary transcripts were elevated in *se* mutants compared to wild type (Fig. 3e and Supplementary Fig. S2), whereas levels of the 21-nt mature transcripts were markedly reduced (Fig. 3f). These observations suggest that SE is required for the processing of pri-miR165 and pri-miR166, and hence for the localized repression of *PHB* gene expression. In dominant *phb-d* mutants, defective miRNA regulation leads to both ectopic *PHB* transcript accumulation and hypomethylation of the *PHB* locus¹¹. Reduction in CG methylation at the *PHB* locus is also detected in *se-3* mutants (Fig. 3g and Supplementary Information),

indicating that *se-3* disrupts *PHB* regulation in a manner similar to *phb-d* alleles.

To test whether SE functions specifically in miRNA-mediated repression of *PHB/PHV* gene expression or is a general regulator of miRNA-based gene regulation, we constructed *se-3; phb-6; phv-5* triple mutants. Reducing HD-Zip III function restored near wild-type morphology to *se-3* mutants. Leaf polarity and rosette size of *se-3; phb-6; phv-5* triple mutants approached those of wild-type plants (Fig. 3h–j), demonstrating that the defects in vegetative development in *se-3* mutants are predominantly a consequence of *PHB* and *PHV* mis-expression. The mild phenotypes that distinguish *se-3; phb-6; phv-5* triple mutants from wild-type plants suggest that SE regulates other genes in addition to *PHB* and *PHV*. These are unlikely to include the most closely related HD-Zip III gene *REVOLUTA (REV)*, the expression of which is unaltered in *se-3*, but could include the more distantly related HD-Zip III gene *CORONA (CNA)*¹⁷, which shows elevated expression in *se-3* plants (Supplementary Fig. S2).

The severe axial patterning phenotypes of *se-3* mutants are attributable to *PHB* mis-expression. Additionally, the weaker *se-1* and *se-2* alleles exhibit phenotypes commonly associated with misregulation of KNOX gene expression, such as the presence of ectopic meristems in *35S::PHB; se-2* plants (Fig. 3b, c) and ectopic stipules in leaves of *asymmetric leaves1; se-1* double mutants¹³. However, *STM*, *BP* and the related gene *KNAT2* are not misregulated in *se* mutants, so SE does not simply act to repress KNOX activity (Fig. 2f, g, data not shown and ref. 13). Rather, SE might limit the competence of shoot tissue to respond to KNOX expression. We tested this hypothesis by broadly expressing an STM–glucocorticoid receptor (STM–GR) fusion protein⁹ in a dose-dependent manner in the *se-2* mutant background. In the presence of 10 nM

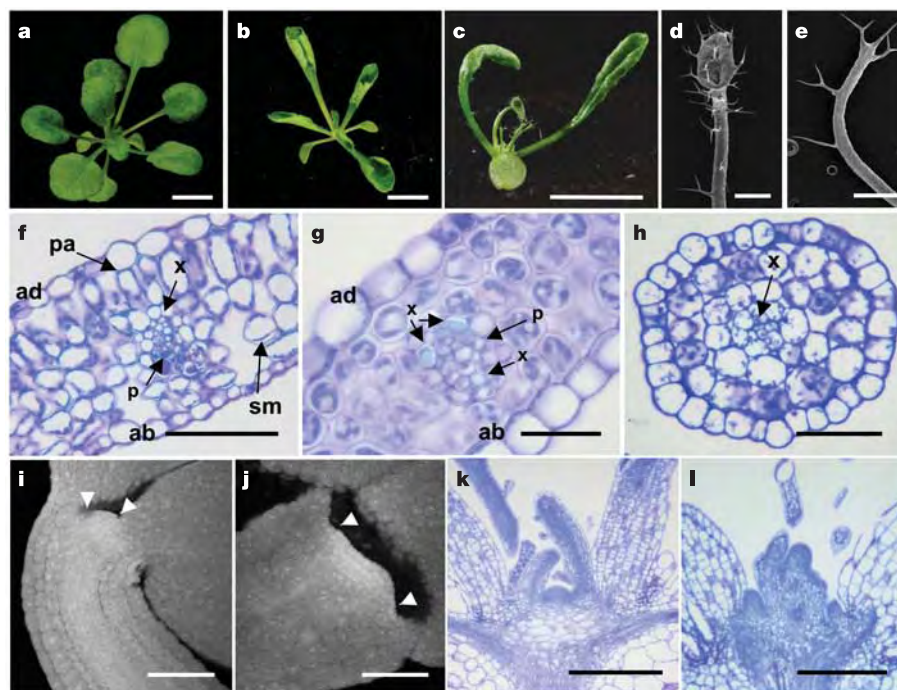


Figure 1 | *se* alleles condition defects in leaf axial patterning and increased meristem size. **a–c**, Rosettes of Columbia (Col, wild type) (**a**), *se-2* (**b**) plants showing upward curling leaves, and *se-3* (**c**) plants showing upward curling and radialized leaves. **d, e**, SEM images of trumpet-shaped (**d**) and radial (**e**) *se-3* leaves. **f–h**, Cross-sections through rosette leaves. **f**, In Col plants, the adaxial side develops a layer of regularly arranged palisade cells, and the abaxial side is characterized by loosely arranged spongy mesophyll cells. Vascular tissue is arranged with xylem on the adaxial and phloem on the abaxial side. **g**, *se-3* plants lack asymmetric cell

differentiation and have ectopic abaxial xylem tissue. **h**, The radially symmetric *se-3* leaf has central xylem tissue and no phloem. **i, j**, Confocal micrographs of mature embryos of Col (**i**), and *se-3* (**j**) showing enlarged SAMs (delineated by arrowheads). **k, l**, Longitudinal sections through Col (**k**) and *se-3* (**l**) seedling apices; with hypertrophic SAMs. Scale bars: 5 mm (**a, b**), 2 mm (**c**), 500 μ m (**d, e**), 10 μ m (**f, g**), 50 μ m (**h–j**), 100 μ m (**k, l**). Abbreviations: ab; abaxial, ad; adaxial p; phloem, pa; palisade, sm; spongy mesophyll, x; xylem.

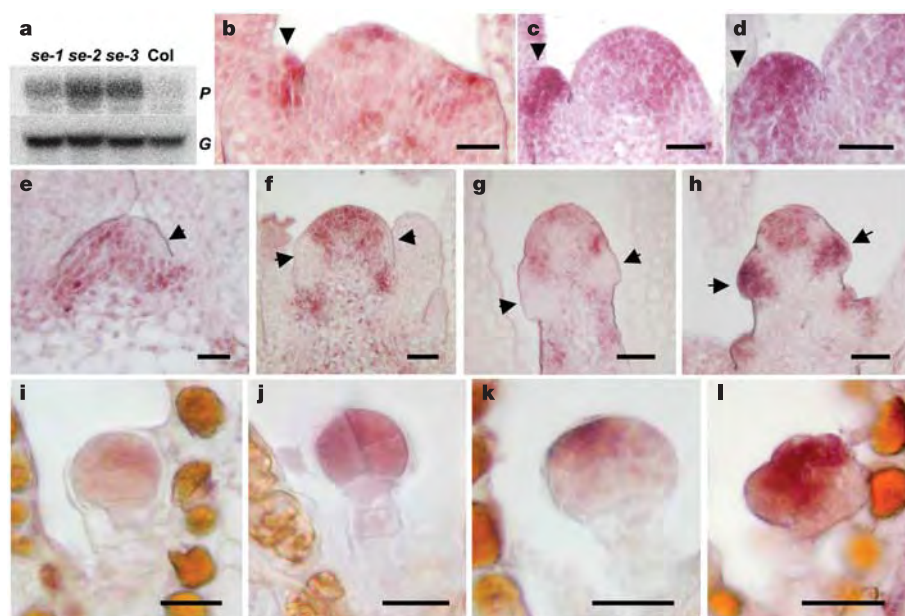


Figure 2 | SE regulates *PHB* expression. **a**, Northern blot analysis of *PHB* (*P*) and *GAPDH* (*G*) transcripts. **b–d**, *In situ* localization of *PHB* mRNA. **b**, In Col apices, *PHB* expression is seen in the centre of the meristem and the adaxial side of leaf primordia (arrowhead). In *se-1* (**c**) and *se-3* (**d**, **h**) apices, expression is expanded both within the meristem and into the abaxial region of leaf primordia. **e–g**, Localization of *STM* mRNA in Col

(**e**) and *se-3* (**f**, **g**) apices. Arrowheads point to areas of *STM* downregulation, which coincide with areas of *PHB* expression (**h**). **i–l**, Col (**i**, **k**) and *se-3* (**j**, **l**) embryos. Elevated and expanded *PHB* expression is evident in both eight-cell and globular stage embryos. Scale bars: 50 μ m (**b–d**), 10 μ m (**e**, **f**, **i–l**), 25 μ m (**g**, **h**).

dexamethasone, *STM-GR* plants showed slightly asymmetric leaves, whereas *STM-GR; se2* plants developed leaves subdivided into lobes (Fig. 4a–d). These leaves had ectopic stipules present in the sinus region between lobes (Fig. 4d), whereas stipules are normally confined to the leaf base. Therefore, the effects of *STM* mis-expression in the leaf are enhanced in *se* mutants, suggesting an antagonistic

relationship between SE and KNOX activity. This antagonism was confirmed by the observation that both *bp* and *stm* mutant phenotypes are suppressed by *se* mutant alleles. *se-1* and *se-2* corrected the propendent silique and shortened pedicel defects¹⁸ observed in *bp-9* mutants (Fig. 4e–l and data not shown), and *se-1* restored organogenic potential to *stm-2* floral meristems (Fig. 4m–o). Therefore,

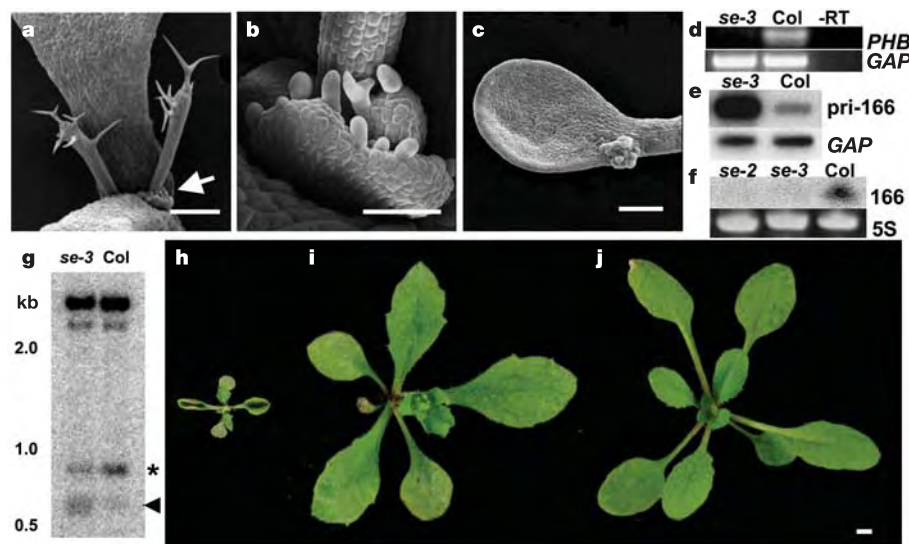


Figure 3 | SE functions in a miRNA gene-silencing pathway. **a–c**, SEM images of *se-2; 35S::PHB*. **a**, Seedling showing radialized leaves. **b**, Magnification of the area denoted by the arrow in **a**, showing organs initiating from an ectopic axillary meristem. **c**, Callus tissue on the adaxial cotyledon surface. **d**, 5' RACE products (~350 bp) generated with nested *PHB*-specific primers, representing messenger RNA cleaved at the miRNA complementary site. *GAPDH* (*GAP*) amplified from the same cDNA samples served as a positive control. **e**, RT-PCR assay of pri-miR166, using *GAPDH* as an amplification control. **f**, Northern blot analysis of miR166, using

ethidium-stained 5S rRNA as a loading control. **g**, Southern blot, for which genomic DNA was digested with methylation-sensitive *HpaII* and probed with a *PHB* probe. Col shows a greater relative intensity of the higher band (asterisk), whereas *se-3* shows greater intensity of the lower band (arrowhead), reflecting hypomethylation of *HpaII* sites. Approximate sizes of a DNA marker are indicated in kb. **h–j**, Rosettes of *se-3; PHB*⁺; *PHV*⁺ (**h**), *se-3; phb-6; phv-5* (**i**) and wild-type F₂ (**j**) plants. Scale bars: 200 μ m (**a**), 50 μ m (**b**), 500 μ m (**c**), 1 cm (**h–j**).

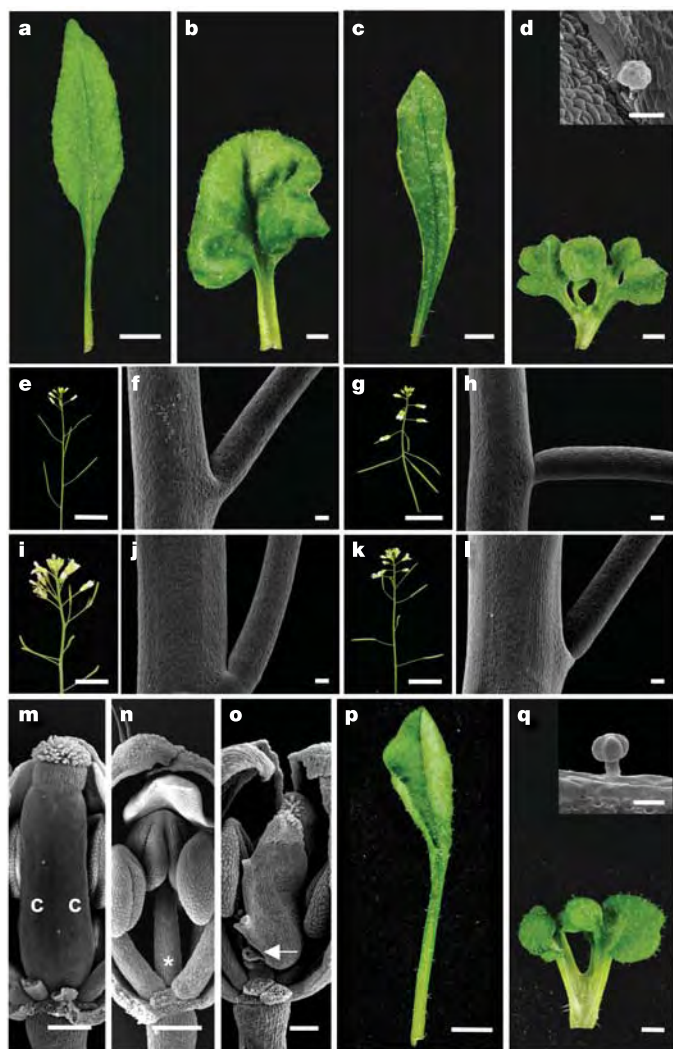


Figure 4 | SE antagonises KNOX activity. **a–d, p, q**, Rosette leaves of plants treated with 10 nM dexamethasone. **a**, Col. **b**, *STM-GR*, showing asymmetric lamina. **c**, *se-2*, showing upward curling. **d**, *STM-GR; se-2*, which is deeply lobed with ectopic stipules (inset). **e–l**, Inflorescences (**e, g, i, k**) and SEM images of the pedicel axil (**f, h, j, l**). **e, f**, Col. **g, h**, *bp-9*, with short, downward-pointing pedicels. **i, j**, *se-2*, which is similar to Col. **k, l**, *bp-9; se-2*, showing rescue of pedicel defects. **m–o**, Developing flowers. **m**, Wild type (*Landsberg erecta*) with two carpels ('c') indistinguishable from *se-1* (not shown). **n**, *stm-2*, lacking carpels in the centre of flower (asterisk). **o**, *se-1; stm-2*, which produces unfused carpels bearing ovules (arrow). **p**, *35S::PHB-1D* transgenic line, with upward curling similar to *se-2*. **q**, *STM-GR; 35S::PHB-1D* is deeply lobed with ectopic stipules (inset). Scale bars: 5 mm (**a, e, g, i, k, p**), 2 mm (**b–d, q**), 100 μ m (**m, n**), 50 μ m (inset in **d**), 20 μ m (inset in **q**).

reduction of SE activity can partially compensate for reduced KNOX function, demonstrating that SE antagonises KNOX action during shoot development.

To assess whether SE-mediated effects on KNOX activity are PHB-dependent, we tested whether a weak *35S::PHB-1D* transgenic line¹⁶ enhances the phenotype of *STM-GR* in a manner comparable to *se*. *STM-GR; se-2* and *STM-GR; 35S::PHB-1D* showed almost indistinguishable phenotypes, suggesting that SE influences KNOX activity through repression of PHB (Fig. 4a–d, p, q). This suggestion is consistent with the findings that *phb-1d* mutants partially suppress *stm*¹⁰, and that both *stm* mutants¹ and triple HD-Zip III gene mutants⁵ are meristemless. Therefore, PHB and KNOX activities converge at common downstream processes that are critical for

meristem development, perhaps through regulation of common target genes (Supplementary Fig. S4 and ref. 19). SE might orchestrate this convergence by delimiting PHB expression.

Our data indicate that SE controls both PHB gene expression and the competence of shoot tissue to respond to KNOX activity. As such, SE defines a regulatory hub that coordinates the activity of pluripotent cells of the SAM with development of their differentiating daughter cells into axially patterned lateral organs^{14,15}. It remains to be seen whether SE activity in regulating cell fate in plants represents the co-option of an ancient eukaryotic function that ascribes specificity to miRNA-based gene regulation, or whether it is a plant-specific innovation.

METHODS

Plant material. *se-2* (SAIL_44_G12) and *se-3* (SALK_083196; ref. 20) are T-DNA insertions in the tenth and first exon, respectively, of the *SERRATE* transcription unit. T-DNA positions were confirmed by polymerase chain reaction (PCR), and mutants were backcrossed twice into the Columbia ecotype. Each insertion line was crossed to *se-1* to confirm allelism. *35S::PHB* and *35S::PHB-1D* plants were generated by transforming plants with constructs pJM2 and pMKB2 (ref. 4), by *Agrobacterium*-mediated floral dipping.

Plant growth conditions. Plants were grown in a greenhouse with supplemental lighting under conditions of 18-h, 20 °C days and 6-h, 16 °C nights, or in a growth cabinet under short-day conditions.

Chemical treatments. Dexamethasone (Sigma) was dissolved in water to a stock concentration of 10^{-6} M, which was added to Murashige-Skoog media to a final concentration of 10^{-9} M.

Histology. Tissue was fixed overnight in 4% paraformaldehyde and 3% glutaraldehyde, embedded in JB4 resin (TAAB), sectioned at 5 μ m, stained with toluidine blue, and viewed on a Leica microscope.

In situ RNA localization. Fixation and hybridization were carried out as previously described¹ using probes for PHB⁴ and STM¹.

Microscopy. Scanning electron microscopy (SEM)²¹ and confocal²² analyses were carried out as previously described.

Reverse transcription PCR gel blot analysis. One microgram total RNA was treated with DNase I and used for complementary DNA synthesis with Superscript RT (Invitrogen). pri-miR166 PCR primers have been described previously⁶. Control reactions containing no cDNA showed no amplification product.

Southern blotting. DNA was extracted from vegetative tissue of three-week-old seedlings and Southern blotting was carried out as previously described¹¹. The probe corresponds to bases 3331–3752 (from ATG) of the PHB genomic locus. The blot was re-probed with a GAPDH probe to confirm complete digestion.

Northern blotting. RNA was extracted from two-week-old seedlings. Sixty micrograms total RNA was passed over a Qiagen RNA column. The flow-through was used for low-molecular-mass analyses, and 10 μ g of eluted RNA was separated on an agarose gel and transferred to a Zeta-Probe nylon membrane (Bio-Rad). The membrane was sequentially hybridized to radiolabelled PHB and GAPDH probes. RNA of low molecular mass was precipitated from the flow-through, separated on a 9% acrylamide, 7 M urea gel, and electroblotted to a Zeta-Probe membrane. The probe was synthesized by α -³²P-ATP end-labelling an oligonucleotide sequence antisense to miR166 (GGGGAATGAAGCCT GGTCCGA), expected to cross-hybridize to miR165. Hybridization was carried out at 40 °C. Blots were replicated using independent RNA samples.

5' RACE. cDNA was generated using a BD-SmartRace kit (BD Biosciences). One microgram of total RNA was used per reaction. PCR amplification with nested PHB-specific primers was carried out at 70 °C annealing temperature. PCR products were separated on a 1.8% agarose gel, and specific products were cloned. Six out of six clones sequenced from Columbia ecotype (Col) showed cleavage in the centre of the miRNA complementary site as shown: (5'-uugggaugaag/ccugguccgg-3'). GAPDH was amplified as a control for cDNA synthesis and amplification. The experiment was replicated using independent RNA samples.

Received 31 March; accepted 15 July 2005.

- Long, J. A., Moan, E. I., Medford, J. I. & Barton, M. K. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of *Arabidopsis*. *Nature* **379**, 66–69 (1996).
- Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K. & Hake, S. A knotted1-like homeobox gene in *Arabidopsis* is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell* **6**, 1859–1876 (1994).

3. Byrne, M. E., Simorowski, J. & Martienssen, R. A. *ASYMMETRIC LEAVES1* reveals *knox* gene redundancy in *Arabidopsis*. *Development* **129**, 1957–1965 (2002).
4. McConnell, J. R. *et al.* Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots. *Nature* **411**, 709–713 (2001).
5. Emery, J. F. *et al.* Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.* **13**, 1768–1774 (2003).
6. Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A. & Timmermans, M. C. microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. *Nature* **428**, 84–88 (2004).
7. Sussex, I. M. Morphogenesis in *Solanum tuberosum* L.: Experimental investigation of leaf dorsiventrality and orientation in the juvenile shoot. *Phytomorphology* **5**, 286–300 (1955).
8. Chuck, G., Lincoln, C. & Hake, S. KNAT1 induces lobed leaves with ectopic meristems when overexpressed in *Arabidopsis*. *Plant Cell* **8**, 1277–1289 (1996).
9. Gallois, J. L., Woodward, C., Reddy, G. V. & Sablowski, R. Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in *Arabidopsis*. *Development* **129**, 3207–3217 (2002).
10. McConnell, J. R. & Barton, M. K. Leaf polarity and meristem formation in *Arabidopsis*. *Development* **125**, 2935–2942 (1998).
11. Bao, N., Lye, K. W. & Barton, M. K. MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev. Cell* **7**, 653–662 (2004).
12. Prigge, M. J. & Wagner, D. R. The *Arabidopsis* *SERRATE* gene encodes a zinc-finger protein required for normal shoot development. *Plant Cell* **13**, 1263–1279 (2001).
13. Ori, N., Eshed, Y., Chuck, G., Bowman, J. L. & Hake, S. Mechanisms that control *knox* gene expression in the *Arabidopsis* shoot. *Development* **127**, 5523–5532 (2000).
14. Englbrecht, C. C., Schoof, H. & Böhm, S. Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. *BMC Genomics* **5**, 39 (2004).
15. Golling, G. *et al.* Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nature Genet.* **31**, 135–140 (2002).
16. Mallory, A. C. *et al.* MicroRNA control of *PHABULOSA* in leaf development: importance of pairing to the microRNA 5' region. *EMBO J.* **23**, 3356–3364 (2004).
17. Green, K. A., Prigge, M. J., Katzman, R. B. & Clark, S. E. CORONA, a member of the class III homeodomain leucine zipper gene family in *Arabidopsis*, regulates stem cell specification and organogenesis. *Plant Cell* **17**, 691–704 (2005).
18. Smith, H. M. & Hake, S. The interaction of two homeobox genes, *BREVIPEDICELLUS* and *PENNYWISE*, regulates internode patterning in the *Arabidopsis* inflorescence. *Plant Cell* **15**, 1717–1727 (2003).
19. Hay, A. *et al.* The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Curr. Biol.* **12**, 1557–1565 (2002).
20. Alonso, J. M. *et al.* Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653–657 (2003).
21. Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* **112**, 1–20 (1991).
22. Bougourd, S., Marrison, J. & Haseloff, J. Technical advance: an aniline blue staining procedure for confocal microscopy and 3D imaging of normal and perturbed cellular phenotypes in mature *Arabidopsis* embryos. *Plant J.* **24**, 543–550 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are indebted to K. Barton and B. Reinhart for providing plasmids 35S::PHB (pJM2) and 35S::PHB-1D (pMKB2) and for sharing data before publication. For helpful discussions and seed stocks we thank N. Ori and S. Hake (*bp-9* seeds) and Y. Eshed and J. Bowman (*phb-6*; *phv-5*; *rev-9/+* seeds). We also thank M. Prigge for *se-1* Ler seeds, R. Sablowski for *STM-GR* seeds and J. Langdale for comments on the manuscript, the Arabidopsis Biological Resource Center and Syngenta for *se-3* and *se-2* seeds respectively, and J. Baker for photography. Finally, we thank A. Bird and I. Henderson for advice on bisulphite sequencing. This work was funded by a BBSRC grant to M.T. S.P.G. is the recipient of a Sainsbury Studentship from the Gatsby Charitable Foundation and A.H. is the recipient of a University of Oxford Glasstone Research Fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.T. (miltos.tsiantis@plants.ox.ac.uk).

Repeated cocaine exposure *in vivo* facilitates LTP induction in midbrain dopamine neurons

Qing-song Liu¹, Lu Pu¹ & Mu-ming Poo¹

Drugs of abuse are known to cause persistent modification of neural circuits, leading to addictive behaviours^{1–5}. Changes in synaptic plasticity in dopamine neurons of the ventral tegmental area (VTA) may contribute to circuit modification induced by many drugs of abuse, including cocaine^{6–13}. Here we report that, following repeated exposure to cocaine *in vivo*, excitatory synapses to rat VTA dopamine neurons become highly susceptible to the induction of long-term potentiation (LTP) by correlated pre- and postsynaptic activity. This facilitated LTP induction is caused by cocaine-induced reduction of GABA_A (γ-aminobutyric acid) receptor-mediated inhibition of these dopamine neurons. In midbrain slices from rats treated with saline or a single dose of cocaine, LTP could not be induced in VTA dopamine neurons unless GABA-mediated inhibition was reduced by bicuculline or picrotoxin. However, LTP became readily inducible in slices from rats treated repeatedly with cocaine; this LTP induction was prevented by enhancing GABA-mediated inhibition using diazepam. Furthermore, repeated cocaine exposure reduced the amplitude of GABA-mediated synaptic currents and increased the probability of spike initiation in VTA dopamine neurons. This cocaine-induced enhancement of synaptic plasticity in the VTA may be important for the formation of drug-associated memory.

To determine the effect of *in vivo* cocaine exposure on synaptic plasticity in VTA dopamine neurons, we examined LTP induction in dopamine neurons in midbrain slices prepared from rats that were given single (one day) or repeated daily intraperitoneal injections (over 5–7 days) of saline or cocaine. The effectiveness of cocaine treatment was shown by the sensitization of locomotor activity (Supplementary Fig. 1). Dopamine neurons were identified by the presence of large I_h currents and distinct firing characteristics^{14,15} (Supplementary Fig. 2). Extracellular stimulation was applied to the rostral region of the VTA, and the evoked excitatory postsynaptic potentials (EPSPs) were monitored by whole-cell recordings from VTA dopamine neurons at -70 mV, close to the reversal potential (-69.7 ± 1.5 mV, $n = 5$) of inhibitory postsynaptic currents (IPSCs). These EPSPs were mediated by the activation of glutamate receptors, as they were completely abolished by the addition of the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor antagonist CNQX (6-cyano-7-nitroquinoxaline-2,3-dione, $20 \mu\text{M}$) and the NMDA (N-methyl-D-aspartate) receptor antagonist AP5 (D-2-amino-5-phosphonopentanoic acid, $50 \mu\text{M}$).

To induce LTP, we used a spike-timing protocol consisting of bursts of EPSP-spike pairs, with the onset of EPSPs preceding the peak of the postsynaptic spike by ~ 5 ms (Fig. 1a; see Methods). This pattern of stimulation was used to mimic bursts of spikes observed in VTA dopamine neurons of behaving rats or monkeys in response to reward-related stimuli^{16,17}. We found that repetitive EPSP-spike pairing induced a long-lasting increase in the amplitude of EPSPs in VTA dopamine neurons in slices obtained from rats treated with

cocaine for 5–7 days (Fig. 1c), but not from rats treated with a single injection of cocaine (one day, Fig. 1e). No LTP was induced in slices from rats that were given time-matched saline injections (Fig. 1b, d). Thus, repeated—but not single—exposure to cocaine facilitated LTP induction in VTA dopamine neurons.

It is known that GABA-mediated inhibition suppresses LTP induction at many excitatory synapses^{18–21}. To determine whether the presence of GABA-mediated inhibition is responsible for the absence of LTP induction in VTA dopamine neurons, we examined LTP induction in the presence of the GABA_A receptor blockers bicuculline or picrotoxin. We found that, in the presence of bicuculline methiodide (BMI, $20 \mu\text{M}$; Fig. 2a, e) or picrotoxin (PTX, $100 \mu\text{M}$; Fig. 2e), EPSP-spike pairing induced robust LTP in VTA dopamine neurons in slices from rats treated with saline for 5–7 days, to a similar extent as that observed for rats treated with cocaine for 5–7 days (Fig. 2b, e). This LTP induction depends on the activation of NMDA receptors, because LTP was absent in the presence of AP5 ($50 \mu\text{M}$, Fig. 2e). Notably, the presence of bicuculline did not further increase the extent of LTP in rats treated with cocaine for 5–7 days (Fig. 2e). Similarly, we found that in the presence of picrotoxin, the paired stimulation induced LTP in dopamine neurons in slices from naive rats, and from rats given one injection of saline or cocaine (Fig. 2c–e). However, application of bicuculline without or immediately after the paired stimulation at -70 mV did not increase the amplitude of EPSPs in rats treated with saline for one day (Supplementary Fig. 3), suggesting that the increase in EPSP amplitude after the paired stimulation was not caused by changes in shunting inhibition but was due to LTP induction. Blocking GABA-mediated inhibition is known to enhance temporal summation of EPSPs²² and NMDA receptor activation during repetitive afferent stimulation²³, and to enhance action potential back-propagation into dendrites²⁴. These mechanisms are likely to be responsible for the facilitatory effect of bicuculline (or picrotoxin) on LTP induction.

To examine whether the increased susceptibility to LTP induction after repeated cocaine exposure is caused by reduced GABA-mediated inhibition, we recorded IPSCs in VTA dopamine neurons in saline- and cocaine-injected rats. Monosynaptic IPSCs were evoked by extracellular stimulation and recorded at $V_c = -20$ mV, in the presence of CNQX ($20 \mu\text{M}$) and AP5 ($50 \mu\text{M}$). The stimulation intensity was gradually increased until the maximal IPSC amplitude was attained¹⁹ (Fig. 3a). The mean amplitude of maximal IPSCs was significantly reduced in slices from cocaine-injected rats compared with IPSCs from saline-injected rats after 5 days of treatment, and this reduction persisted for at least 12 days after withdrawal from the 7-day cocaine injection (Fig. 3a). The change in IPSCs was not due to differences in the stimulation intensity (data not shown). The passive membrane properties of these dopamine neurons also remained unchanged after 5–7 days of cocaine treatment (Supplementary Table 1). Furthermore, the cocaine-induced reduction in IPSC amplitude was cell-specific, because the average maximal IPSC

¹Division of Neurobiology, Department of Molecular and Cell Biology, Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720, USA.

amplitude in non-dopamine neurons was not changed (Fig. 3b). The cocaine-induced reduction in IPSC amplitude is of postsynaptic origin, because local puffing of the GABA_A receptor agonist muscimol (20 μ M) to the soma of dopamine neurons produced smaller currents (Fig. 3c), and the amplitude of miniature IPSCs (mIPSCs) recorded was smaller for 5–7-day cocaine-treated rats than for saline-treated rats, but the mean frequency of mIPSCs was not significantly different (Fig. 3d). Together, these findings indicate that repeated cocaine exposure selectively reduces postsynaptic GABA-responsiveness in dopamine neurons.

Consistent with reduced inhibition of VTA dopamine neurons, we found that the probability of action potential initiation in response to presynaptic stimuli was increased after repeated cocaine exposure. We constructed input–output curves for action potentials by varying the stimulation intensity and measuring the probability of action potential firing using cell-attached patch recording. The slope of the input–output curve for rats treated with cocaine for 5–7 days was significantly higher than that for rats treated with saline for the same period of time. Furthermore, the slopes for both groups increased and converged to the same level after bath addition of bicuculline (20 μ M, Fig. 3e and Supplementary Table 2). Similarly, we examined the input–output relationship for evoked excitatory postsynaptic

currents (EPSCs) by voltage-clamping the dopamine neuron at the mean resting membrane potential ($V_c = -54$ mV). The slope of the input–output curve for EPSCs was significantly higher in cocaine-treated rats than that in 5-day saline-treated rats, and again, bicuculline increased the slope for both groups (Supplementary Fig. 4). Thus, repeated cocaine exposure increases spike output primarily through the reduction of GABA-mediated inhibition, but potential changes in intrinsic excitability might also contribute. Such hyperexcitability might endow VTA dopamine neurons with an additional capacity to promote reward-dependent learning, in addition to the enhanced LTP induction described above.

To further test the idea that reduced GABA-mediated inhibition is responsible for the facilitation of LTP induction in dopamine neurons, we examined LTP induction in slices from 5–7-day saline-treated rats in the presence of different concentrations of bicuculline. Consistent with findings that LTP induction is highly sensitive to the level of GABA-mediated inhibition^{19–21,25}, we found that LTP was

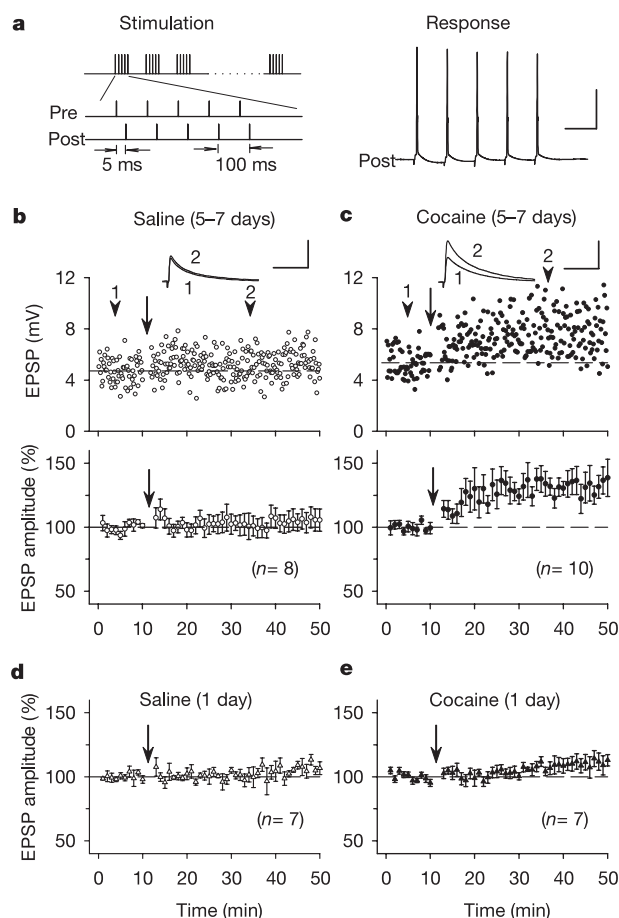


Figure 1 | Repeated cocaine exposure facilitates LTP induction in VTA dopamine neurons. **a**, Left, the protocol for LTP induction consists of 20 bursts of EPSP–spike pairs. Right, a typical postsynaptic response during one burst of paired stimuli. Scale bar, 30 mV, 100 ms. **b**, **c**, Examples (top) and summary (bottom) of normalized EPSP amplitude before and after paired stimulation for rats treated with saline (**b**) or cocaine (**c**) for 5–7 days. Arrows indicate application of the LTP induction protocol. Scale bars, 5 mV, 50 ms. **d**, **e**, The effect of paired stimulation on EPSP amplitude for rats given a single dose of saline or cocaine (1 day). Error bars indicate s.e.m.

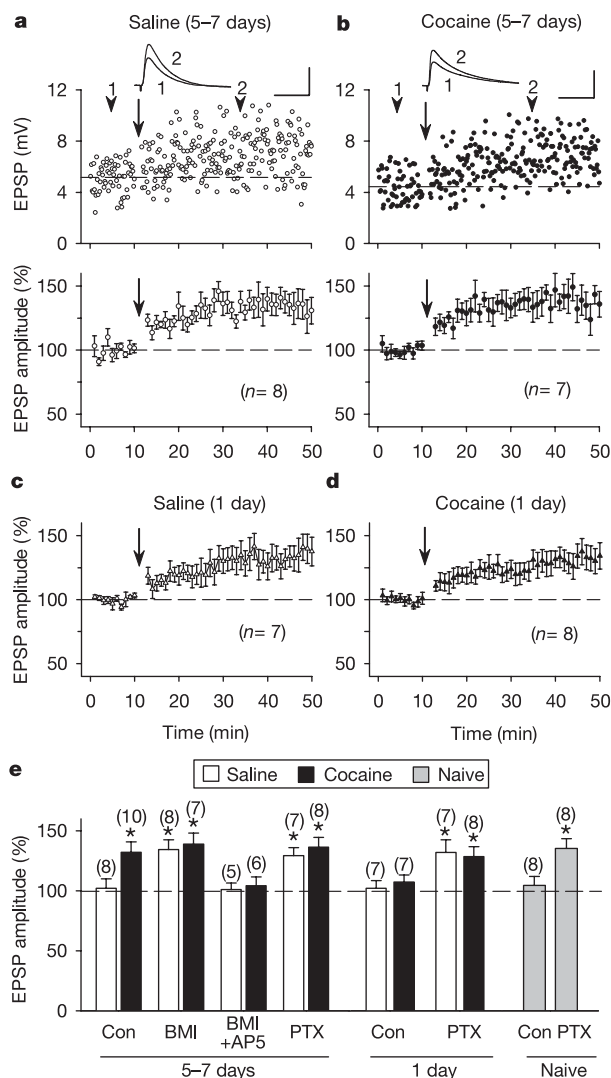


Figure 2 | The effect of blocking GABA-mediated inhibition on LTP induction in dopamine neurons. **a–d**, In the presence of 20 μ M bicuculline (BMI) (**a**, **b**) or 100 μ M picrotoxin (PTX) (**c**, **d**), paired stimulation induced LTP for all saline- or cocaine-treated rats. Data are presented in the same manner as those in Fig. 1. Scale bars, 5 mV, 50 ms. **e**, Summary of LTP induction under various conditions. Control (con), without any blockers. The total number of neurons examined is shown in parentheses. Data significantly different from the baseline value of the same experiment are marked by an asterisk ($P < 0.01$, paired t -test). Error bars indicate s.e.m.

induced in the presence of $1\ \mu\text{M}$ but not $0.3\ \mu\text{M}$ bicuculline (Fig. 4a, c). Notably, more complete blockade of GABA-mediated inhibition using $20\ \mu\text{M}$ bicuculline did not further increase the extent of potentiation (Fig. 4c and Supplementary Fig. 5a, b). Furthermore, application of diazepam, a benzodiazepine agonist known to enhance the activation of GABA_A receptors²⁶, prevented LTP induction in 5–7-day cocaine-treated rats at a concentration of $1\ \mu\text{M}$, but not at $0.1\ \mu\text{M}$. The effect of diazepam ($1\ \mu\text{M}$) on LTP induction was abolished by bicuculline (Fig. 4b, c). The effectiveness of diazepam in enhancing the activation of GABA_A receptors was shown by the dose-dependent increase in amplitude and the decay time constant of IPSCs (Supplementary Fig. 5c–f). Together, these results indicate that GABA-mediated inhibition normally suppresses LTP induction in VTA dopamine neurons, and repeated cocaine exposure facilitates LTP induction by reducing GABA-mediated inhibition to below a critical level.

Previous studies have shown that a single exposure to cocaine *in vivo* results in an increase in the ratio of AMPA- to NMDA-receptor-mediated EPSCs in VTA dopamine neurons in midbrain slices^{8,9}, and that repeated cocaine exposure produces no further increase in this ratio¹⁰. Furthermore, a single exposure to cocaine occluded LTP induction in these neurons *in vitro* in the presence of picrotoxin, using an induction protocol of pairing presynaptic stimulation with steady postsynaptic depolarization⁸. These results suggest that a single exposure to cocaine produces saturated potentiation of excitatory inputs onto dopamine neurons. We note,

however, that in the presence of picrotoxin, this induction protocol was not efficient for LTP induction even in the saline-treated group⁸ (Fig. 5a), but LTP could be induced readily in slices from rats exposed to single or repeated cocaine treatment using the spike-timing induction protocol. Thus, the discrepancy in LTP induction may be attributed in part to the induction protocol used.

In the present study, we confirmed that the AMPA-to-NMDA receptor ratio was increased after a single exposure to cocaine (Fig. 5b, c). The finding that the cocaine-induced increase in the AMPA-to-NMDA receptor ratio did not preclude LTP induction indicates that either only a subset of synapses are potentiated *in vivo* by cocaine exposure, or alternatively that LTP at these synapses is expressed by mechanisms independent of the increased AMPA-to-NMDA receptor ratio (for example, an increase in presynaptic transmitter release^{27,28}). Regardless of the mechanism, our finding suggests that the strength of excitatory synaptic inputs to dopamine neurons still remains within a modifiable range after repeated exposure to cocaine, and that drug-associated stimuli can exert a persistent imprint on VTA circuits.

By inducing LTP in VTA dopamine neurons in the absence of GABA_A receptor blockers, we were able to uncover the cocaine-induced reduction in GABA-mediated inhibition and the facilitatory effect of cocaine on LTP induction. This reduced inhibition might help to transform drug-associated stimuli into incentive stimuli, by allowing such stimuli to induce persistent synaptic modification in VTA dopamine neurons. Such incentive sensitization may underlie

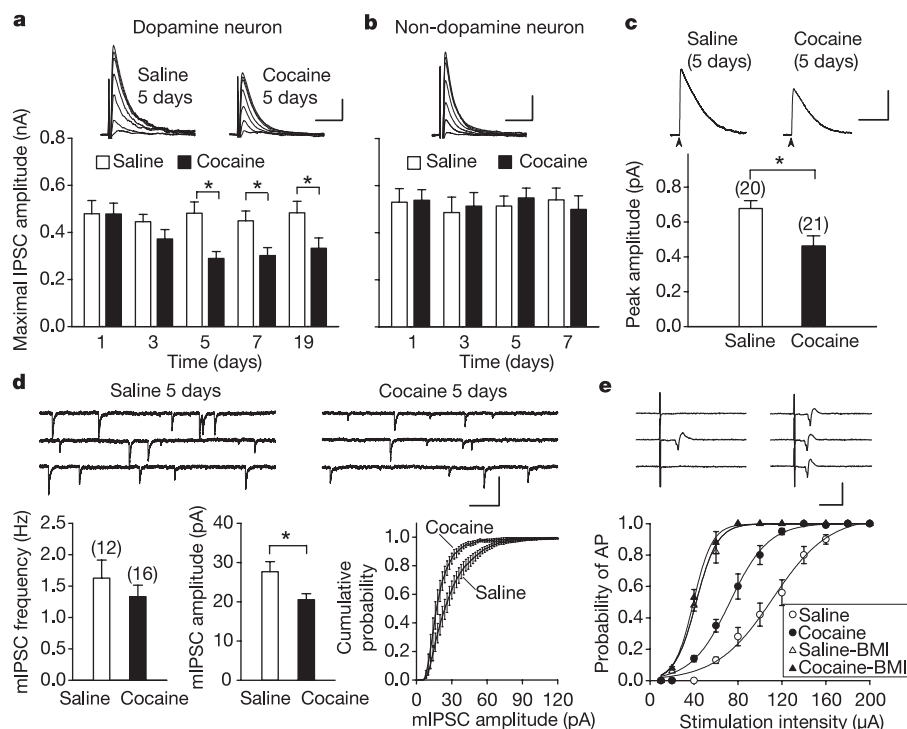


Figure 3 | Effects of cocaine exposure on GABA-mediated inhibition of VTA neurons. **a**, Repeated cocaine treatment reduces the amplitude of maximal IPSCs in dopamine neurons. Top, sample IPSCs in response to stimuli of incremental intensities (20–140 μA). Scale bar, 100 pA, 15 ms. Bottom, the mean amplitude of maximal IPSCs after 1–7 days of saline or cocaine treatment, and 12 days after withdrawal from 7 days of cocaine treatment (19 days); $n = 17$ –25 for each bar. **b**, Mean maximal IPSC amplitudes in non-dopamine neurons ($n = 15$ –19). Scale bar, 150 pA, 15 ms. **c**, Samples (top) and average amplitudes (bottom) of currents induced by puffing muscimol ($20\ \mu\text{M}$, arrowhead) at the soma of dopamine neurons after 5–7 days of saline or cocaine treatment. Asterisk, $P < 0.01$ (t -test). Scale bar, 500 pA, 3 s. **d**, Repeated cocaine treatment (5–7 days) had no effect on mean

mIPSC frequency but decreased mean mIPSC amplitude (asterisk, $P < 0.01$, t -test) and significantly shifted the mIPSC amplitude distribution (bin size 2 pA) towards smaller values ($P < 0.001$, Kolmogorov–Smirnov test). Sample traces of mIPSCs are shown at the top. Scale bar, 80 pA, 250 ms. **e**, Effect of repeated cocaine exposure on the probability of spike initiation in dopamine neurons. Top, examples of action potentials induced by $80\ \mu\text{A}$ stimulation for a saline-treated rat (5 days) before (left) and after (right) BMI ($20\ \mu\text{M}$). Scale bar, 30 pA, 10 ms. Bottom, input–output curves (fitted with the Boltzmann function) of the probability of action potentials induced by graded stimuli for rats treated with saline or cocaine for 5–7 days. Error bars indicate s.e.m. ($n = 8$ –10).

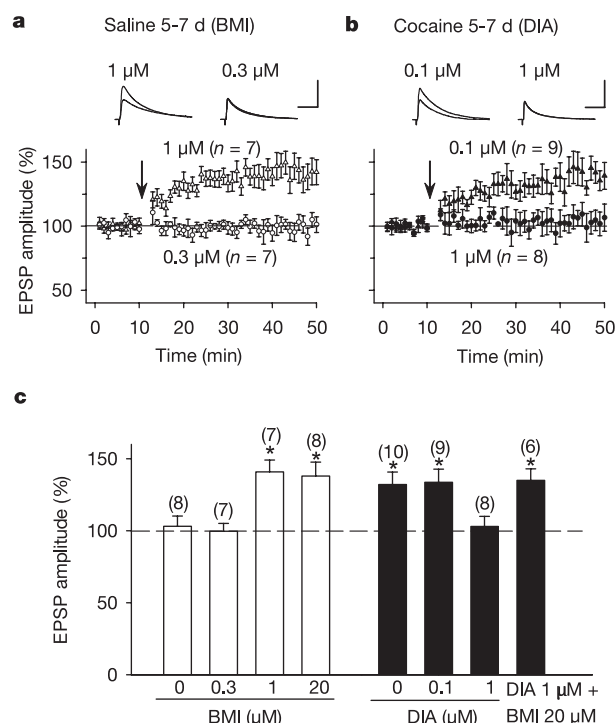


Figure 4 | A critical level of GABA-mediated inhibition regulates LTP induction in dopamine neurons. **a**, The effects of partial blocking of GABA_A receptors with bicuculline (BMI, 0.3 and 1 μM) on LTP induction in rats treated with saline for 5–7 days. Scale bar, 5 mV, 50 ms. **b**, The effects of enhancing GABA_A receptor activation with diazepam (DIA, 0.1 and 1 μM) on LTP induction in rats treated with cocaine for 5–7 days. Scale bar, 5 mV, 50 ms. **c**, Summary graph showing concentration-dependent effects of bicuculline and diazepam on LTP induction in rats treated with saline (open bars) or cocaine (filled bars) for 5–7 days. Asterisk, $P < 0.01$ (paired t -test). The total number of neurons examined is shown in parentheses. Error bars indicate s.e.m.

drug craving and compulsive drug-seeking behaviour³. On the other hand, we note that the reduced inhibition and *in vivo* synaptic potentiation are unlikely to be responsible for locomotor sensitization, because locomotor sensitization precedes the reduction in inhibition (compare Fig. 3a and Supplementary Fig. 1), and there is no correlation between the increase in the AMPAR-to-NMDA receptor ratio and locomotor sensitization^{10,11}. It has been reported that γ -vinyl-GABA, which enhances GABA-mediated inhibition,

reduces cocaine-seeking behaviour in rats²⁹ and cocaine craving/taking in cocaine addicts in clinical trials³⁰, but the underlying cellular mechanism is largely unknown. Our findings suggest that preventing drug-induced synaptic plasticity through enhancement of GABA-mediated inhibition might in part account for the effectiveness of γ -vinyl-GABA in reducing cocaine addiction.

METHODS

Cocaine treatment and slice preparation. Male Sprague-Dawley rats (21–40 days old) were given daily intraperitoneal injections with either saline (0.9% NaCl, 1 ml kg⁻¹) or cocaine (15 mg kg⁻¹ dissolved in saline) for 1–7 days. Rats were anaesthetized by halothane and killed by decapitation 20–24 h or 12 days (for Fig. 3a only) after the last injection, in accordance with institutional guidelines. Horizontal midbrain slices (250-μm thick) were cut using a vibratome (Vibratome Company). Slices were prepared at 4–6 °C in a solution containing 110 mM choline chloride, 2.5 mM KCl, 1.25 mM NaH₂PO₄, 0.5 mM CaCl₂, 7 mM MgSO₄, 26 mM NaHCO₃, 25 mM glucose, 11.6 mM sodium ascorbate and 3.1 mM sodium pyruvate. The slices were incubated in artificial cerebrospinal fluid (ACSF) containing 125 mM NaCl, 3 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂, 1.25 mM NaH₂PO₄, 26 mM NaHCO₃ and 10 mM glucose. The solutions were saturated with 95% O₂ and 5% CO₂.

Electrophysiology. Whole-cell or cell-attached recordings were made using a patch clamp amplifier (Multiclamp 700B, Axon Instruments) under infrared-DIC (differential interference contrast) microscopy. Data acquisition and analysis were performed using a digitizer (DigiData 1322A, Axon Instruments) and the analysis software pClamp 9 (Axon Instruments) and Mini Analysis 6 (Synaptosoft). Signals were filtered at 2 Hz and sampled at 10 Hz. For presynaptic stimulation, a bipolar tungsten stimulation electrode (WPI) placed ~150 μm rostral to the recording electrode was used to stimulate synaptic inputs, using a stimulation pulse of duration 40 μs and frequency 0.1 Hz. For recording EPSPs, neurons were current-clamped at -70 mV. For recording IPSCs, neurons were voltage-clamped at -20 mV in the presence of CNQX (20 μM) and AP5 (50 μM). For recording of EPSC–IPSC sequences, neurons were clamped near the resting membrane potential (-54 mV). For recording muscimol-induced currents, muscimol (20 μM) was ejected through a glass micropipette (tip resistance 8–10 MΩ) with a picospritzer (100 ms, 30 psi) to the soma of the recorded dopamine neuron ($V_c = -20$ mV), with the pipette tip located ~20 μm from the soma surface. The whole-cell recording pipette (tip resistance 3–5 MΩ) was filled with a solution containing 140 mM potassium gluconate, 5 mM KCl, 10 mM HEPES, 0.2 mM EGTA, 2 mM MgCl₂, 4 mM MgATP, 0.3 mM Na₂GTP and 10 mM Na₂-phosphocreatine, pH 7.2 (with KOH). For recording evoked IPSCs and muscimol-induced currents, the experimenters were blind to the treatment history of the rat. For recording mIPSCs, potassium gluconate was replaced with KCl (140 mM) in the internal solution, and TTX (0.5 μM), CNQX (10 μM) and AP5 (25 μM) were added to the bath. Series resistance (15–30 MΩ) or input resistance (300–500 MΩ) was monitored throughout the whole-cell recording and data were discarded if the resistance changed by more than 20%. All recordings were performed at 33 ± 1 °C with the help of an automatic temperature controller (Warner Instruments).

The spike-timing protocol for LTP induction consisted of 20 bursts of

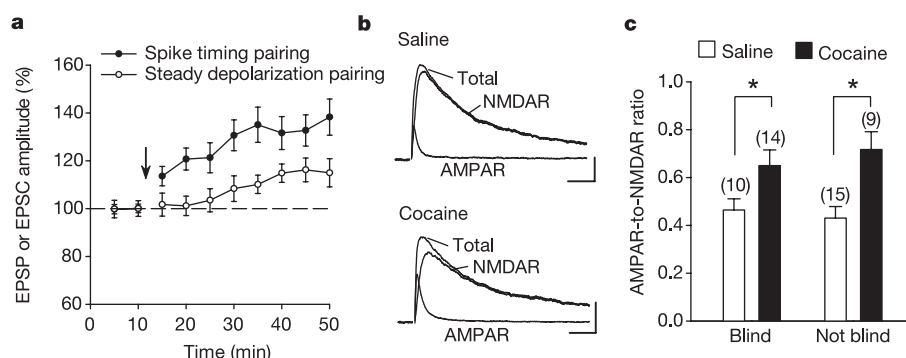


Figure 5 | Comparison of LTP induced by two different protocols and cocaine-induced changes in the AMPA-to-NMDA receptor ratio. **a**, Effects of spike-timing and steady-depolarization paired protocols on the amplitude of EPSPs (spike-timing, same data as Fig. 2c but averaged at 5-min intervals) and EPSCs (steady depolarization) for rats treated with saline for one day ($n = 7$). **b**, Examples of total EPSCs, AMPA-receptor- (AMPA) and

NMDA-receptor- (NMDAR)-mediated EPSCs in dopamine neurons from rats treated with saline or cocaine for one day (see Methods). Scale bars, 50 pA, 15 ms. **c**, Summary of mean AMPAR-to-NMDAR ratios obtained by the experimenters, who were either blind or not blind to the treatment history of the rats. The total number of neurons examined is shown in parentheses. Asterisk, $P < 0.05$ (t -test). Error bars indicate s.e.m.

EPSP-spike pairs, with each burst consisting of 5 paired stimuli delivered at 100-ms intervals (10 Hz) and an inter-burst interval of about 5 s. The post-synaptic spikes were evoked by injection of depolarizing current pulses (1–2 nA, 2–3 ms), with the onset of EPSPs preceding the peak of postsynaptic spikes by ~5 ms. Evoked EPSPs were sampled at 0.1 Hz before and after LTP induction. For quantifying the magnitude of LTP, 60 consecutive EPSPs were averaged 10 min before and 20 min after the end of the induction protocol. Data are given as mean $\pm$ s.e.m. For data comparison, paired or unpaired Student's *t*-tests and one-way analysis of variance (ANOVA) were used. The steady-depolarization pairing protocol for LTP induction consists of 200 stimuli of synaptic inputs (1 Hz) paired with a steady postsynaptic depolarization (+10 mV), as described previously⁸.

For measurements of the AMPA-to-NMDA receptor ratio, the dopamine neuron was voltage-clamped at +40 mV. First, a stable baseline recording of total EPSCs was obtained. The NMDA receptor antagonist AP5 (50 μ M) was then applied to the bath for 5–10 min to isolate fast AMPA-receptor-mediated EPSCs. Digital subtraction of AMPA-receptor-EPSCs from the total EPSCs from the same neuron yielded NMDA-receptor-EPSCs. An average of 12–20 EPSCs were collected for each type of EPSC. The bath solution contained picrotoxin (100 μ M). The intracellular solution contained 120 mM CsCH₃SO₃, 20 mM HEPES, 0.4 mM EGTA, 2.8 mM NaCl, 5 mM TEA-Cl, 2 mM MgCl₂, 2.5 mM MgATP and 0.3 mM GTP, pH 7.2 (with CsOH).

Received 31 March; accepted 22 July 2005.

- Nestler, E. J. Molecular basis of long-term plasticity underlying addiction. *Nature Rev. Neurosci.* **2**, 119–128 (2001).
- Yao, W. D. *et al.* Identification of PSD-95 as a regulator of dopamine-mediated synaptic and behavioural plasticity. *Neuron* **41**, 625–638 (2004).
- Robinson, T. E. & Berridge, K. C. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* **18**, 247–291 (1993).
- Thompson, A. M., Gosnell, B. A. & Wagner, J. J. Enhancement of long-term potentiation in the rat hippocampus following cocaine exposure. *Neuropharmacology* **42**, 1039–1042 (2002).
- Robinson, T. E. & Kolb, B. Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology* **47** (suppl.), 33–46 (2004).
- Kauer, J. A. Learning mechanisms in addiction: synaptic plasticity in the ventral tegmental area as a result of exposure to drugs of abuse. *Annu. Rev. Physiol.* **66**, 447–475 (2004).
- Faleiro, L. J., Jones, S. & Kauer, J. A. Rapid synaptic plasticity of glutamatergic synapses on dopamine neurons in the ventral tegmental area in response to acute amphetamine injection. *Neuropsychopharmacology* **29**, 2115–2125 (2004).
- Ungless, M. A., Whistler, J. L., Malenka, R. C. & Bonci, A. Single cocaine exposure *in vivo* induces long-term potentiation in dopamine neurons. *Nature* **411**, 583–587 (2001).
- Saal, D., Dong, Y., Bonci, A. & Malenka, R. C. Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron* **37**, 577–582 (2003).
- Borgland, S. L., Malenka, R. C. & Bonci, A. Acute and chronic cocaine-induced potentiation of synaptic strength in the ventral tegmental area: electrophysiological and behavioural correlates in individual rats. *J. Neurosci.* **24**, 7482–7490 (2004).
- Dong, Y. *et al.* Cocaine-induced potentiation of synaptic strength in dopamine neurons: behavioural correlates in GluRA(–/–) mice. *Proc. Natl Acad. Sci. USA* **101**, 14282–14287 (2004).
- Wolf, M. E., Sun, X., Mangiavacchi, S. & Chao, S. Z. Psychomotor stimulants and neuronal plasticity. *Neuropharmacology* **47** (suppl.), 61–79 (2004).
- Kelley, A. E. Memory and addiction: shared neural circuitry and molecular mechanisms. *Neuron* **44**, 161–179 (2004).
- Johnson, S. W. & North, R. A. Two types of neurone in the rat ventral tegmental area and their synaptic inputs. *J. Physiol. (Lond.)* **450**, 455–468 (1992).
- Jones, S. & Kauer, J. A. Amphetamine depresses excitatory synaptic transmission via serotonin receptors in the ventral tegmental area. *J. Neurosci.* **19**, 9780–9787 (1999).
- Hyland, B. I., Reynolds, J. N., Hay, J., Perk, C. G. & Miller, R. Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience* **114**, 475–492 (2002).
- Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).
- Wigstrom, H. & Gustafsson, B. Facilitated induction of hippocampal long-lasting potentiation during blockade of inhibition. *Nature* **301**, 603–604 (1983).
- Huang, Z. J. *et al.* BDNF regulates the maturation of inhibition and the critical period of plasticity in mouse visual cortex. *Cell* **98**, 739–755 (1999).
- Bissiere, S., Humeau, Y. & Luthi, A. Dopamine gates LTP induction in lateral amygdala by suppressing feedforward inhibition. *Nature Neurosci.* **6**, 587–592 (2003).
- Meredith, R. M., Floyer-Lea, A. M. & Paulsen, O. Maturation of long-term potentiation induction rules in rodent hippocampus: role of GABAergic inhibition. *J. Neurosci.* **23**, 11142–11146 (2003).
- Pouille, F. & Scanziani, M. Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science* **293**, 1159–1163 (2001).
- Herron, C. E., Williamson, R. & Collingridge, G. L. A selective *N*-methyl-D-aspartate antagonist depresses epileptiform activity in rat hippocampal slices. *Neurosci. Lett.* **61**, 255–260 (1985).
- Larkum, M. E., Zhu, J. J. & Sakmann, B. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* **398**, 338–341 (1999).
- Artola, A., Brocher, S. & Singer, W. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* **347**, 69–72 (1990).
- Eghbali, M., Curmi, J. P., Birnir, B. & Gage, P. W. Hippocampal GABA_A channel conductance increased by diazepam. *Nature* **388**, 71–75 (1997).
- Stevens, C. F. & Wang, Y. Changes in reliability of synaptic function as a mechanism for plasticity. *Nature* **371**, 704–707 (1994).
- Bolshakov, V. Y. & Siegelbaum, S. A. Regulation of hippocampal transmitter release during development and long-term potentiation. *Science* **269**, 1730–1734 (1995).
- Gardner, E. L. *et al.* Gamma-vinyl GABA, an irreversible inhibitor of GABA transaminase, alters the acquisition and expression of cocaine-induced sensitization in male rats. *Synapse* **46**, 240–250 (2002).
- Brodie, J. D., Figueroa, E. & Dewey, S. L. Treating cocaine addiction: from preclinical to clinical trial experience with gamma-vinyl GABA. *Synapse* **50**, 261–265 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by grants from the US National Institutes of Health.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.-m.P. (mpoo@berkeley.edu) or Q.-s.L. (qslu@berkeley.edu).

LETTERS

A network-based analysis of systemic inflammation in humans

Steve E. Calvano^{1*}, Wenzhong Xiao^{2*}, Daniel R. Richards³, Ramon M. Felciano³, Henry V. Baker^{4,5}, Raymond J. Cho³, Richard O. Chen³, Bernard H. Brownstein⁶, J. Perren Cobb⁶, S. Kevin Tschoeke⁵, Carol Miller-Graziano⁷, Lyle L. Moldawer⁵, Michael N. Mindrinos², Ronald W. Davis², Ronald G. Tompkins⁸, Stephen F. Lowry¹ & the Inflammation and Host Response to Injury Large Scale Collaborative Research Program†

Oligonucleotide and complementary DNA microarrays are being used to subclassify histologically similar tumours, monitor disease progress, and individualize treatment regimens^{1–5}. However, extracting new biological insight from high-throughput genomic studies of human diseases is a challenge, limited by difficulties in recognizing and evaluating relevant biological processes from huge quantities of experimental data. Here we present a structured network knowledge-base approach to analyse genome-wide transcriptional responses in the context of known functional interrelationships among proteins, small molecules and phenotypes. This approach was used to analyse changes in blood leukocyte gene expression patterns in human subjects receiving an inflammatory stimulus (bacterial endotoxin). We explore the known genome-wide interaction network to identify significant functional modules perturbed in response to this stimulus. Our analysis reveals that the human blood leukocyte response to acute systemic inflammation includes the transient dysregulation of leukocyte bioenergetics and modulation of translational machinery. These findings provide insight into the regulation of global leukocyte activities as they relate to innate immune system tolerance and increased susceptibility to infection in humans.

Inflammation is a hallmark of many human diseases^{6–8}. We focus on blood leukocytes and other tissues of critically injured patients, in order to better elucidate the mechanisms underlying systemic inflammatory responses⁹. This approach cannot be fully replicated using animal models or human cell lines, and studies of injury in humans can be complicated by antecedent illnesses and concurrent treatment regimes that may alter the recovery process. To our knowledge, no study has evaluated the genome-wide response to systemic inflammation in the context of a fully predictable recovery. Here we combine genome-wide expression analysis with a new bioinformatics method to identify functional networks responsible for the systemic activation and spontaneous resolution of a well-defined inflammatory challenge.

Gene expression in whole blood leukocytes was determined immediately before and at 2, 4, 6, 9 and 24 h after the intravenous administration of bacterial endotoxin to four healthy human subjects. Four additional subjects were studied under identical conditions but without endotoxin administration. The infusion of

endotoxin activates innate immune responses and presents with physiological responses of brief duration¹⁰. Notably, there is an initial proinflammatory phase and a subsequent counterregulatory phase, with resolution of virtually all clinical perturbations within 24 h.

K-means cluster and principal component analyses were first used to visualize the overall response to endotoxin administration. Figure 1a reveals probe sets clustered by K-mean analysis, where each bin has a distinct endotoxin-induced temporal pattern. The signal intensity of 5,093 probe sets—representing 3,714 unique genes—out of a total of >44,000 probe sets changed significantly in response to endotoxin, whereas no significant changes were observed in control subjects (estimated false discovery rate <0.1%). Of the 5,093 probe sets identified, over half showed reduced abundance at 2, 4, 6 and 9 h, returning to baseline by 24 h (see bins 0–4). In contrast, a smaller number of probe sets were induced by 2 h (bins 5, 6), and the remaining probe sets showed a delayed response, peaking at 4–9 h but returning to baseline by 24 h (bins 7–9).

Cluster and principal component analyses describe overall changes in apparent gene expression, but provide few insights into the biological processes and signalling networks invoked in propagation and resolution of the inflammatory response. Identifying the perturbed biological networks underlying this complex clinical phenotype requires systematic analysis in the context of known mammalian biology, derived from basic and clinical research.

Using a web-based entry tool developed by Ingenuity Systems Inc., findings presented in peer-reviewed scientific publications were systematically encoded into an ontology by content and modelling experts. Using over 200,000 full-text scientific articles, a knowledge base of more than 9,800 human, 7,900 mouse and 5,000 rat genes was manually curated and supplemented with curated relationships parsed from MEDLINE abstracts. A molecular network of direct physical, transcriptional and enzymatic interactions observed between mammalian orthologues—the observed ‘interactome’—was computed from this knowledge base. The resulting network contains molecular relationships involving over 8,000 orthologues with a high degree of connectivity. On average, individual genes have 11.5 interaction partners (median 4.0), of which 7.2 represent direct physical interactions (median 3.0). Every gene interaction in the network is supported by published information. For example, the

¹Department of Surgery, UMDNJ-Robert Wood Johnson Medical School, New Brunswick, New Jersey 08903, USA. ²Stanford Genome Technology Center, Palo Alto, California 94304, USA. ³Ingenuity Systems Inc, Mountain View, California 94043, USA. ⁴Departments of Molecular Genetics and Microbiology, and ⁵Department of Surgery, University of Florida College of Medicine, Gainesville, Florida 32610, USA. ⁶Department of Surgery, Washington University in St Louis, St Louis, Missouri 63110, USA. ⁷Department of Surgery, University of Rochester School of Medicine, Rochester, New York 14642, USA. ⁸Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.

*These authors contributed equally to this work.

†Lists of participants and affiliations appear at the end of the paper

immediate neighbourhood for the *RELA* gene (NF- κ B p65, Supplementary Fig. 1) includes 150 genes and 619 direct interactions, derived from 7,118 findings curated from 847 published articles.

The observed interactome provides a framework for structuring the existing knowledge regarding mammalian biology, and enables a new analytical approach that objectively examines experimental data in the context of known genome-wide interactions in order to identify significant functional modules. This method is applicable to data of high-throughput platforms such as microarray expression profiling, polymorphism analysis and proteomics. Furthermore, the

original literature detailing the genetic interactions can be accessed to further examine and verify the findings.

For a better understanding of the temporal response of gene expression in the innate immune system, we constructed a prototypical inflammatory cell containing 292 representative genes and detailing all direct interactions in our database (Fig. 2). Closer inspection of the temporal response reveals the fine structure of dynamic changes in RNA abundance by highlighting the transient and self-limiting nature of this response. As an example, the apparent expression of several secreted proinflammatory cytokines and chemokines (*TNFSF2* (*TNF*), *IL1A*, *IL1B*, *CXCL1* (*GRO α*), *CXCL2* (*GRO β*), *CCL2* (*MCP-1*), *CXCL8* (*IL-8*) and *CXCL10*) reached a maximum 2–4 h after endotoxin administration, consistent with early activation of innate immunity. Subsequently, the expression of several members of the nuclear factor kappa/relA family of transcription factors (*NFKB1*, *NFKB2*, *RELA* and *RELB*) reached their zenith.

The time period 4–6 h after endotoxin injection seemed critical, as the expression of a number of transcription factors was increased, including both those that initiate and those that limit the innate immune response. In the former group, these included the signal transducer and activators of transcription (*STAT* genes), and the cAMP-response element-binding protein (*CREB*) and CCAAT/enhancer binding protein (*CEBP*) gene families. Transcription factors limiting the innate immune response included suppressor of cytokine signalling 3 (*SOCS3*) and *IKBK* genes. There was also a delay (4–6 h) in increased mRNA abundance of secreted and membrane-associated proteins that limit the inflammatory response, including *IL1RAP*, *IL1R2*, *IL10* and *TNFRSF1A*. Together, these data comprehensively document the temporal modulation of genes controlling the innate immune response in a human model that progresses from an acute proinflammatory phase to unencumbered counter-regulation, concluding with full recovery and a normal phenotype.

To further elucidate the global changes during inflammation and subsequent return to homeostasis, we sought to computationally decipher the principal networks involved. The specificity of connections for each gene was calculated, as defined by the percentage of its direct connections to other genes showing significant transcriptional changes. A network pathway was initiated by the gene with the highest specificity of connections, and was propagated according to the descent of the specificity. Individual significant pathways identified by a statistical likelihood calculation ($P < 0.0001$) were merged to represent the biological processes.

Our global representation of the inflammatory response to endotoxin, shown in Fig. 3a, comprises a network of 1,556 genes and their interactions. This network consists of a subset of 1,214 genes (78%) responsive to *in vivo* endotoxin administration, and 342 additional, highly interconnected genes. The gestalt of the temporal response to endotoxin is suggested at the level of interconnecting functional modules, which could not be readily extracted from the experimental data alone (Fig. 1). Simultaneous survey and evaluation of the sub-network regions enables us to identify new endotoxin-responsive modules in addition to the innate immunity network described above. Examples of the diversity in such modules include (a) increased expression of components of the superoxide-producing phagocyte NADPH-oxidase system, a multicomponent enzyme important for host defence¹¹, (b) decreased expression of the major histocompatibility (MHC) II complex, consistent with reduced antigen presentation following endotoxin stimulation^{12,13}, (c) decreased expression of the TCP1 ring complex required for folding of cytoskeletal proteins, (d) increased expression in the family of tubulin-A microtubule genes, (e) suppressed expression of several subunits of the anaphase-promoting complex, which has a key role in cell-cycle regulation, and (f) reduced integrin- α and - β chain expression, affecting cell-cell and cell-matrix adhesion (see Supplementary Methods 2, containing Supplementary Fig. 3a–f).

Further, significant decreases in messenger RNA abundance were

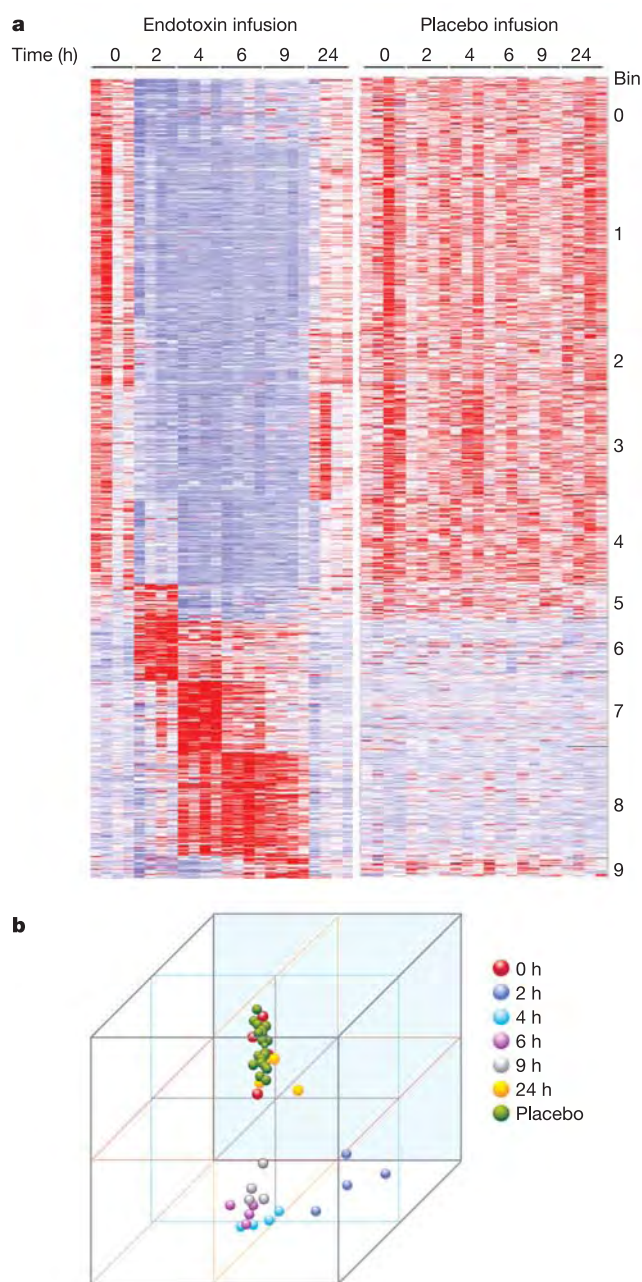
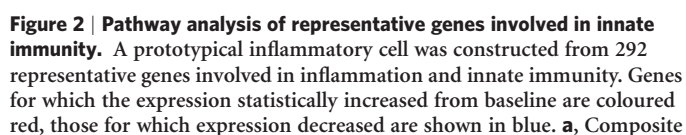


Figure 1 | Gene expression profiles of circulating leukocytes in response to bacterial endotoxin infusion. Samples from eight healthy volunteers were tested at baseline (0 h) and 2, 4, 6, 9 and 24 h after intravenous administration of endotoxin (four subjects) or vehicle (four subjects). **a**, Significant (false discovery rate of $<0.1\%$) probe sets (5,093) were subjected to *K*-means clustering into ten bins (0–9). Probe sets for which the abundance was above the mean are shown in red, below the mean are shown in blue, and equivalent to the mean are in white. **b**, Principal component plot of the significant probe sets at the indicated times after endotoxin administration.



changes in apparent expression over 24 h, identifying nodes and interactions. **b**, Temporal changes in apparent expression. The response to endotoxin administration in blood leukocytes can be viewed as an integrated cell-wide response, propagating and resolving over time.

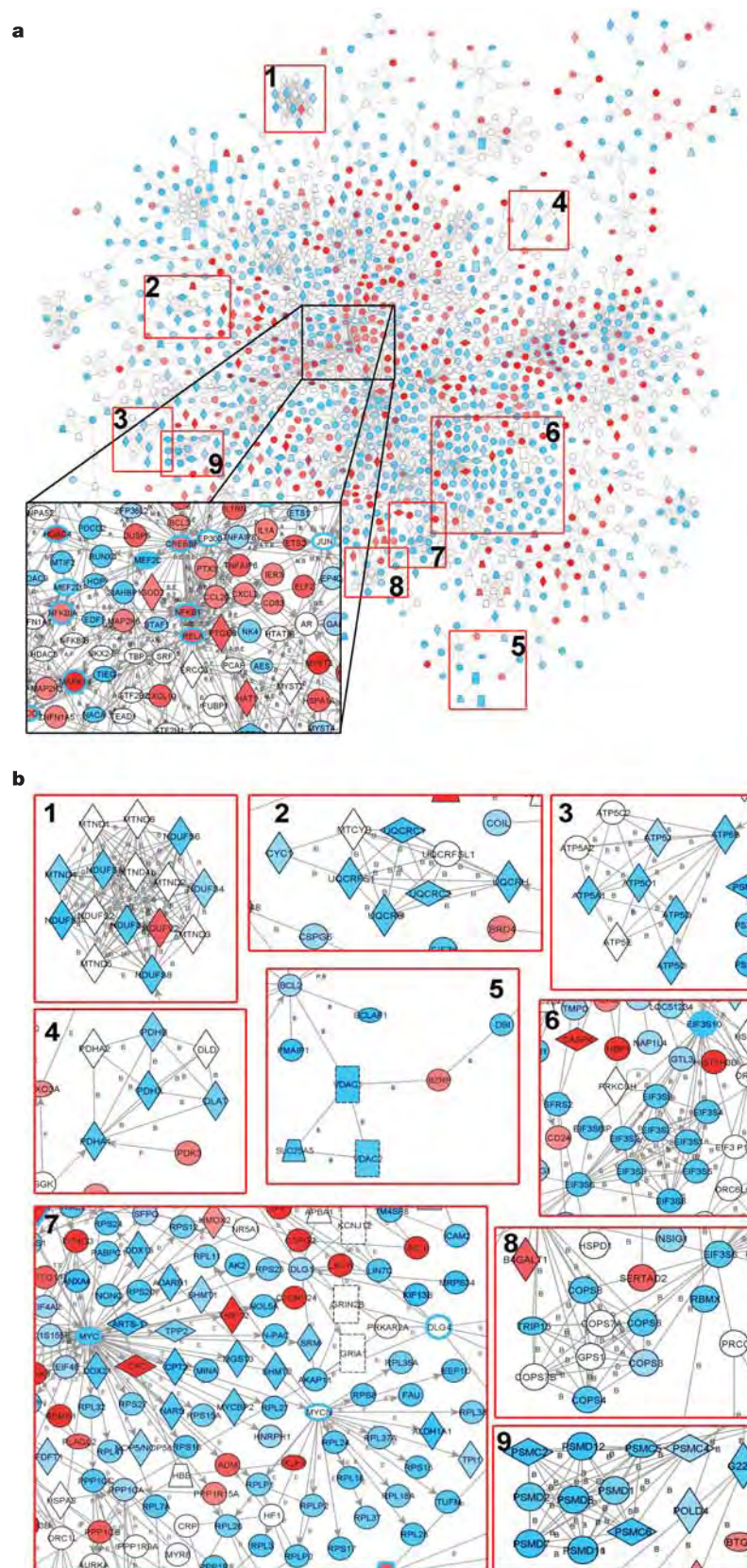


Figure 3 | Network representation of the biological processes underlying the temporal response of blood leukocytes to *in vivo* endotoxin administration. **a**, The network consists of 1,214 genes showing perturbed expression, and 342 genes highly interconnected to this group (red, increased; blue, decreased expression). **b**, Selected regions of the network, highlighting several groups of genes. Group 1, mitochondrial respiratory chain complex I (*NDUF* genes). Group 2, mitochondrial respiratory chain

complex III (*UQCRC* genes). Group 3, ATP synthase complex (*ATP5* genes). Group 4, pyruvate dehydrogenase complex. Group 5, mitochondrial permeability transition pore complex. Group 6, elongation initiation factor complex (*EIF3* genes). Group 7, ribosomal proteins (*RPL*, *RPS* genes). Group 8, COP9 signalosome (*COPS* genes). Group 9, proteasome (*PSM* genes).

observed in the mitochondrial respiratory chain complexes I–V (*NDUF* in Fig. 3b (Group 1), *UQC*R (Group 2) and ATP synthase genes (Group 3)). Gene expression was also decreased in the pyruvate dehydrogenase complex (PDH, Fig. 3b, Group 4), which generates, via acetyl-CoA and the tricarboxylic acid (TCA) cycle, reduced coenzymes required for ATP synthesis during mitochondrial oxidative phosphorylation. A concomitant increase in expression of pyruvate dehydrogenase kinase-3 (*PKD3*), an inhibitor of PDH, was observed. Expression of the voltage dependent anion channel (VDAC) and adenine nucleotide translocator (*SLC25A5*), components of the mitochondrial permeability transition pore (MPTP) complex, were decreased to similar extents, whereas expression of the antagonistic benzodiazepine receptor (*BZRP*) was increased (Fig. 3b, Group 5). MPTP activation has previously been considered an early event in apoptosis, leading to mitochondrial membrane depolarization and release of cytochrome *c*. However, recent reports have indicated a primary role for MPTP activation in oxidative-stress- and calcium-overload-induced necrotic cell death^{14,15}. Thus, reduction in transcripts for MPTP components is consistent with a protective response to the oxidative stress associated with endotoxin challenge.

As active secretory cells, leukocytes devote a substantial amount of energy expenditure to protein synthesis¹⁶. In concert with the suppression of modules participating in energy production, expression was decreased for the elongation initiation factor complex (EIF3 in Fig. 3b, Group 6), a large number of ribosomal proteins (*RPS*, *RPL* genes in Fig. 3b, Group 7), the RNA polymerase II complex, and also in the functional modules of the ATP/ubiquitin-dependent protein degradation pathway, the COP9 signalosome (Fig. 3b, Group 8) and the proteasome (*PSM* genes in Fig. 3b, Group 9).

Here we use a knowledge-based network analysis to reveal concerted dysregulation of functional modules in mitochondrial bioenergetics, protein synthesis and protein degradation in human blood leukocytes during an abbreviated and self-limiting episode of inflammation. These findings document a reprioritization of the transcriptional regulatory programme in leukocytes in response to endotoxin. Furthermore, the marked suppression of these important functional networks suggests that leukocytes exposed to inflammatory stimuli may have an altered capacity to sustain subsequent immune challenges, as observed during innate immune system tolerance.

Human disease phenotypes are manifested by the malfunctioning of multiple functional modules interrelated in physiological regulatory systems. The overwhelming diversity of possible genome-wide interactions and gene expression patterns limit effective learning from experimental data alone. Network analyses using comprehensive knowledge of mammalian biology can greatly reduce the hypothesis space, enabling identification of new functional modules perturbed in the disease process. Here we demonstrate that, upon acute systemic inflammation, the human blood leukocyte response includes widespread suppression at the transcriptional level of mitochondrial energy production and protein synthesis machinery. A decrease in high-energy substrates has been observed in the muscle and liver of critically ill patients, and in animal models of sepsis, burn injury and endotoxemia^{17–19}. In addition, direct disruption of mitochondrial complexes by cellular-stress-derived mediators (for example, reactive oxygen species) has been suggested for necrotic cell death in animal models of sepsis²⁰ and reperfusion injury^{14,15}.

The erosion in functional networks identified above represents a normal adaptive process aimed at re-establishing homeostasis, and yet might contribute to global leukocyte defects—such as tolerance and increased susceptibility to infection—observed in critically injured patients. These perturbations in functional modules are at the level of mRNA transcripts, and will require subsequent confirmation at the protein level. Further identification of the specific cell populations showing these changes in gene expression will require the isolation and enrichment of specific leukocyte subpopulations. Finally, it will be important to confirm whether patients manifesting

systemic inflammation show similar perturbations of the functional modules identified here as do otherwise healthy, endotoxin-challenged subjects.

METHODS

Human endotoxin model. Eight healthy male and female subjects between 18 and 40 years of age provided written informed consent. Details of endotoxin administration to human subjects have been summarized elsewhere^{10,21}. Subjects were intravenously administered either NIH Clinical Center Reference Endotoxin, (CC-RE-Lot 2) at a dose of 2 ng kg⁻¹ body weight (*n* = 4, one female and three males) or 0.9% sodium chloride (*n* = 4, one female and three males) over a 5-min period. Blood samples were collected before endotoxin infusion (0 h) and 2, 4, 6, 9 and 24 h after infusion.

Blood sampling. Blood was collected and lysis buffer (bicarbonate-buffered ammonium chloride solution, 0.826% NH₄Cl, 0.1% KHC₃, 0.0037% Na₄EDTA in H₂O) was added at a ratio of 20:1 (lysis buffer:blood). Samples were then incubated at room temperature until erythrocyte lysis was complete (~5–7 min). Leukocytes were recovered by centrifugation (400 g, 4 °C) and washed once in ice-cold phosphate buffered saline. Leukocyte pellets were then resuspended in 8 ml RLT buffer (Qiagen) and the samples sheared ten times with an 18-gauge needle attached to a 10-ml syringe. Samples were then immediately frozen and kept at –70 °C until RNA extraction was required.

Leukocyte RNA isolation. Total cellular RNA was isolated from the leukocyte pellets using a commercial kit (RNeasy, Qiagen). Purity was confirmed by spectrophotometry (*A*₂₆₀/*A*₂₈₀ ratio) and capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent Inc).

cRNA synthesis and chip hybridization. cRNA synthesis was performed using 4 µg total cellular RNA, hybridized onto Hu133A and Hu133B oligonucleotide arrays (Affymetrix), and processed according to the protocol outlined by Affymetrix, with a few modifications.

Microarray data analysis. A total of 44,924 probe sets on the Hu133A and Hu133B arrays were analysed. Normalization was performed using dChip²², and expression level was modelled using the perfect match only model. Probe sets identified as absent on all arrays (using MicroArray Suite v5, Affymetrix) were not included in further analysis. Probe sets significantly perturbed after bacterial endotoxin administration were identified using significance analysis of microarrays (SAM) (multiclass response)²³, with an estimated false discovery rate of <0.1% on the basis of 1,000 permutations. The resulting 5,093 probe sets were subjected to *K*-means clustering to ten bins using Cluster and TreeView²⁴, and principal component analysis. The same analysis was applied to control data, but no temporal changes of the probe sets reached the significant level (false discovery rate <0.1%). A total of 3,714 unique genes were identified from 4,895 probe sets with mapped Entrez GeneIDs (<http://www.ncbi.nih.gov/Entrez/>). More information is at <http://www.gluegrant.org/>.

Verification of genes showing significant transcriptional changes. An additional six healthy subjects (one female and five males) were administered 2 ng kg⁻¹ (body weight) endotoxin, and blood samples were collected before (0 h) and after (2 and 6 h) endotoxin infusion. At either 2 or 6 h, ~88% of the significant probe sets in the initial study were again identified as significant (false discovery rate <1%). See Supplementary Methods 1.

Development of the observed interactome of mammalian genes. The Ingenuity Pathways Knowledge Base (KB) is the largest curated database of previously published findings on mammalian biology from the public literature (Ingenuity Systems). Reports on individual studies of genes in human, mouse or rat were first identified from peer-reviewed publications, and findings were then encoded into an ontology by content and modelling experts. Manual extraction and curation probably results in more specific and comprehensive interactions, with far fewer false-positives than automated alternatives (for example, natural language processing and high-throughput screening).

Network analysis using the knowledge base was used to further identify direct interactions between mammalian orthologues. For example, the knowledge of functional interactions (for example, phosphorylation) was combined with knowledge of protein functions (for example, kinase activity) to algorithmically infer a direct biochemical interaction. Every resulting gene interaction has supporting literature findings available online. More information is in Supplementary Methods 2.

Identification of significant pathways in biological processes. The following steps were used: (1) Genes identified as significant from the experimental data sets were overlaid onto the interactome. Focus genes were identified as the subset having direct interaction(s) with other genes in the database. (2) The specificity of connections for each focus gene was calculated by the percentage of its connections to other significant genes. The initiation and growth of pathways proceeded from genes with the highest specificity of connections. Each pathway

had a maximum of 35 genes. (3) Pathways of highly interconnected genes were identified by statistical likelihood using the following equation:

$$\text{Score} = -\log_{10} \left(1 - \sum_{i=0}^{f-1} \frac{C(G, i)C(N - G, s - i)}{C(N, s)} \right)$$

where N is the number of genes in the genomic network, of which G are focus genes, for a pathway of s genes, f of which are focus genes. $C(n, k)$ is the binomial coefficient. (4) Pathways with a score greater than 4 ($P < 0.0001$) were combined to form a composite network representing the underlying biology of the process.

Received 17 May; accepted 4 July 2005.

Published online 31 August 2005.

- Heller, R. A. *et al.* Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl Acad. Sci. USA* **94**, 2150–2155 (1997).
- Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
- Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).
- Bittner, M. *et al.* Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**, 536–540 (2000).
- Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
- Coussens, L. M. & Werb, Z. Inflammation and cancer. *Nature* **420**, 860–867 (2002).
- Ridker, P. M., Hennekens, C. H., Buring, J. E. & Rifai, N. C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. *N. Engl. J. Med.* **342**, 836–843 (2000).
- Cheung, C. Y. *et al.* Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: a mechanism for the unusual severity of human disease? *Lancet* **360**, 1831–1837 (2002).
- Bone, R. C. *et al.* Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest* **101**, 1644–1655 (1992).
- Fong, Y. M. *et al.* The acute splanchnic and peripheral tissue metabolic response to endotoxin in humans. *J. Clin. Invest.* **85**, 1896–1904 (1990).
- Dinauer, M. C., Orkin, S. H., Brown, R., Jesaitis, A. J. & Parkos, C. A. The glycoprotein encoded by the X-linked chronic granulomatous disease locus is a component of the neutrophil cytochrome *b* complex. *Nature* **327**, 717–720 (1987).
- Wolk, K., Docke, W. D., von Baehr, V., Volk, H. D. & Sabat, R. Impaired antigen presentation by human monocytes during endotoxin tolerance. *Blood* **96**, 218–223 (2000).
- Wolk, K., Kunz, S., Crompton, N. E., Volk, H. D. & Sabat, R. Multiple mechanisms of reduced major histocompatibility complex class II expression in endotoxin tolerance. *J. Biol. Chem.* **278**, 18030–18036 (2003).
- Nakagawa, T. *et al.* Cyclophilin D-dependent mitochondrial permeability transition regulates some necrotic but not apoptotic cell death. *Nature* **434**, 652–658 (2005).
- Baines, C. P. *et al.* Loss of cyclophilin D reveals a critical role for mitochondrial permeability transition in cell death. *Nature* **434**, 658–662 (2005).
- Buttgereit, F. & Brand, M. D. A hierarchy of ATP-consuming processes in mammalian cells. *Biochem. J.* **312**, 163–167 (1995).
- Crouser, E. D., Julian, M. W., Blaho, D. V. & Pfeiffer, D. R. Endotoxin-induced mitochondrial damage correlates with impaired respiratory activity. *Crit. Care Med.* **30**, 276–284 (2002).
- Padfield, K. E. *et al.* Burn injury causes mitochondrial dysfunction in skeletal muscle. *Proc. Natl Acad. Sci. USA* **102**, 5368–5373 (2005).
- Brealey, D. *et al.* Association between mitochondrial dysfunction and severity and outcome of septic shock. *Lancet* **360**, 219–223 (2002).
- Fink, M. P. Cytopathic hypoxia. Mitochondrial dysfunction as a mechanism contributing to organ dysfunction in sepsis. *Crit. Care Clin.* **17**, 219–237 (2001).
- Kalil, A. C. *et al.* Effects of drotrecogin alfa (activated) in human endotoxemia. *Shock* **21**, 222–229 (2004).
- Li, C. & Wong, W. H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA* **98**, 31–36 (2001).
- Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98**, 5116–5121 (2001).
- Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14863–14868 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. M. Coyle for clinical assistance, and J. Wilhelmy, A. Kumar, S. MacMillan and A. Abouhamze for technical assistance. This work was supported by an Injury and the Host Response to Inflammation Large Scale Collaborative Research Program Award from the National Institute of General Medical Sciences (to R.G.T.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to R.W.D. (dbowe@stanford.edu).

Inflammation and Host Response to Injury Large Scale Collaborative Research Program Paul E. Bankey¹, Timothy R. Billiar², David G. Camp³, George Casella⁴, Irshad H. Chaudry⁵, Mashkoo A. Choudhry⁵, Charles Cooper²⁶, Asit De¹, Constance Elson⁶, Bradley Freeman⁷, Richard L. Gamelli⁸, Celeste Campbell-Finnerty⁹, Nicole S. Gibran¹⁰, Douglas L. Hayden⁶, Brian G. Harbrecht², David N. Herndon⁹, Jureta W. Horton¹¹, William J. Hubbard⁵, John L. Hunt¹², Jeffrey Johnson¹³, Matthew B. Klein¹⁴, James A. Lederer¹⁵, Tanya Logvinenko⁶, Ronald V. Maier¹⁰, John A. Mannick¹⁵, Philip H. Mason²⁶, Bruce A. McKinley¹⁶, Joseph P. Minei¹¹, Ernest E. Moore¹³, Frederick A. Moore¹⁶, Avery B. Nathens¹⁰, Grant E. O'Keefe¹⁰, Laurence G. Rahme¹⁷, Daniel G. Remick¹⁸, David A. Schoenfeld⁶, Martin Schwacha⁵, Michael B. Shapiro¹⁹, Geoffrey M. Silver⁸, Richard D. Smith³, John D. Storey²⁰, Mehmet Toner²¹, H. Shaw Warren²², Michael A. West¹⁹

¹⁰Affiliations for participants: Department of Surgery, University of Rochester School of Medicine, Rochester, New York 14642, USA. ¹¹Department of Surgery, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213, USA. ¹²Pacific Northwest National Laboratory, Richland, Washington 99352, USA. ¹³Department of Statistics, University of Florida, Gainesville, Florida 32611, USA. ¹⁴Department of Surgery, University of Alabama School of Medicine, Birmingham, Alabama 35294, USA. ¹⁵Department of Biostatistics, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ¹⁶Department of Surgery, Washington University School of Medicine, St Louis, Missouri 63110, USA. ¹⁷Department of Surgery, Loyola University Stritch School of Medicine, Maywood, Illinois 60153, USA. ¹⁸Department of Surgery, University of Texas Medical Branch, Shriners Burns Hospital, Galveston, Texas 77550, USA. ¹⁹Department of Surgery, University of Washington Harborview Medical Center, Seattle, Washington 98104, USA. ²⁰Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ²¹Division of Trauma, Burns, and Critical Care, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ²²Department of Surgery, University of Colorado Denver Health Medical Center, Denver, Colorado 80204, USA. ²³Burn Center and Division of Plastic Surgery, University of Washington Harborview Medical Center, Seattle, Washington 98104, USA. ²⁴Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ²⁵Department of Surgery, University of Texas Houston Health Science Center, Houston Medical School, Houston, Texas 77030, USA. ²⁶Department of Molecular Biology, Massachusetts General Hospital Harvard Medical School, Boston, Massachusetts 02114, USA. ²⁷Department of Medical Science, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. ²⁸Department of Surgery, Northwestern University Medical School, Chicago, Illinois 60611, USA. ²⁹Department of Biostatistics, University of Washington, Seattle, Washington 98195, USA. ³⁰Center for Engineering in Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ³¹Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.

LETTERS

Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines

Qinghua Shi^{1†} & Randall W. King¹

Although mutations in cell cycle regulators or spindle proteins can perturb chromosome segregation^{1–7}, the causes and consequences of spontaneous mitotic chromosome nondisjunction in human cells are not well understood. It has been assumed that nondisjunction of a chromosome during mitosis will yield two aneuploid daughter cells. Here we show that chromosome nondisjunction is tightly coupled to regulation of cytokinesis in human cell lines, such that nondisjunction results in the formation of tetraploid rather than aneuploid cells. We observed that spontaneously arising binucleated cells exhibited chromosome mis-segregation rates up to 166-fold higher than the overall mitotic population. Long-term imaging experiments indicated that most binucleated cells arose through a bipolar mitosis followed by regression of the cleavage furrow hours later. Nondisjunction occurred with high frequency in cells that became binucleated by furrow regression, but not in cells that completed cytokinesis to form two mononucleated cells. Our findings indicate that nondisjunction does not directly yield aneuploid cells, but rather tetraploid cells that may subsequently become aneuploid through further division. The coupling of spontaneous segregation errors to furrow regression provides a potential explanation for the prevalence of hyperdiploid chromosome number and centrosome amplification observed in many cancers^{8,9}.

Despite pathways such as the spindle checkpoint that promote accurate chromosome segregation^{10,11}, nondisjunction still occurs with significant frequency in human cells^{12,13}. Because chromosome mis-segregation is irreversible, aneuploid cells will accumulate unless mechanisms exist to eliminate them or suppress their proliferation. To identify such mechanisms, we asked whether the accuracy of chromosome segregation is compromised in cells that show altered behaviour in subsequent steps of cell division, such as cytokinesis. We therefore compared the frequency of chromosome mis-segregation in all mitotic cells with the frequency measured in spontaneously arising binucleated cells. We analysed anaphase or telophase cells containing bipolar spindles by fluorescence *in situ* hybridization (FISH) using centromeric probes for chromosomes 8 and 12 (Fig. 1a; see also Supplementary Fig. 1). In diploid human keratinocytes immortalized with telomerase (N/TERT-1; ref. 14), we observed mis-segregation of chromosome 8 and 12 in 0.05% and 0.10% of mitotic cells, respectively (Fig. 1c), comparable to reported rates in human lymphocytes and fibroblasts^{12,13}. HeLa cells, which are aneuploid but maintain a stable karyotype¹⁵, showed higher rates of mis-segregation, 0.24% per mitosis for chromosome 8 and 0.39% for chromosome 12 (Fig. 1d). Similar results were obtained in immortalized prostate epithelial (PrEC) cells¹⁶ (see Supplementary Fig. 1).

The frequency of chromosome mis-segregation was found to be much higher in spontaneously arising binucleated cells (Fig. 1b; see also Supplementary Fig. 1). In binucleated N/TERT-1 cells, 8.3% of

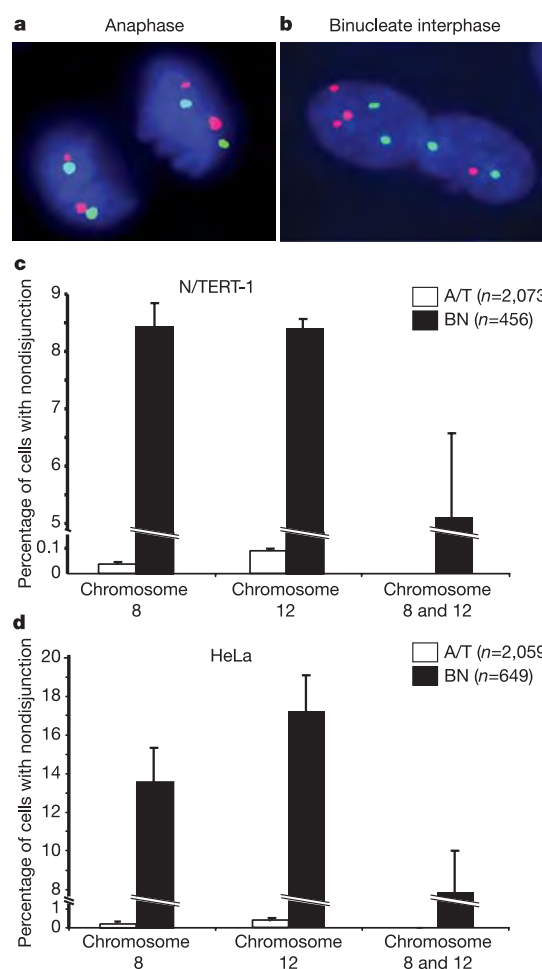


Figure 1 | Chromosome mis-segregation occurs at high frequency in spontaneously arising binucleated cells. **a, b**, Examples of mitotic (**a**) and binucleated interphase (**b**) N/TERT-1 cells analysed by fluorescence *in situ* hybridization using chromosome 8-(red) and 12-(green) specific centromeric probes. Example of normal segregation (**a**) and chromosome 8 nondisjunction in a binucleated cell (**b**). **c**, Quantification of chromosome nondisjunction in binucleated (BN) and anaphase/telophase (A/T) N/TERT-1 cells with normal bipolar spindles. **d**, Quantification of chromosome mis-segregation in HeLa cells (for images see Supplementary Fig. 1). Error bars are one standard deviation, from results of 4 (**c**) or 5 (**d**) independent experiments.

¹Department of Cell Biology, Harvard Medical School, 240 Longwood Ave, Boston, Massachusetts 02115, USA. [†]Present address: Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230026, China.

cells mis-segregated chromosome 8 or 12 during the preceding mitosis, increases of 166-fold and 83-fold, respectively, compared to the frequencies measured in mitotic cells. In binucleated HeLa cells, the frequency of cells with chromosome mis-segregation was 13.7% for chromosome 8 and 17.4% for chromosome 12, increases of 57-fold and 45-fold, respectively. In binucleated PrEC cells, non-disjunction rates were 14.5% for chromosome 8 and 15.0% for chromosome 12, increases of 48-fold and 50-fold, respectively (see Supplementary Fig. 1). In all cell lines, a fraction of binucleated cells mis-segregated multiple chromosomes. Together these data indicate that spontaneously arising binucleated cells have substantially elevated chromosome mis-segregation rates compared to the overall mitotic population.

Our data suggested that cytokinesis might be inhibited in cells that spontaneously mis-segregate chromosomes during the preceding mitosis. To determine whether treatments that induce chromosome mis-segregation also increase the frequency of binucleation, we treated HeLa cells with short interfering RNAs (siRNAs) or drugs that would reduce the accuracy of chromosome segregation¹⁰. Reducing expression of the spindle checkpoint protein Mad2 or the protease separase by siRNA increased the rate of chromosome mis-segregation and doubled the frequency of binucleated cells (see Supplementary Fig. 2). Partial inhibition of topoisomerase II with low doses of ICRF-193 (ref. 17) also increased chromosome mis-

segregation and binucleation. Similar results were obtained with VP-16 and nocodazole (see Supplementary Fig. 3), supporting the possibility that chromosome segregation errors are functionally coupled to the regulation of cytokinesis.

To determine how spontaneous binucleation occurs, we imaged HeLa and N/TERT-1 cells expressing green fluorescent protein (GFP)-tagged histone 2B (ref. 18). Expression of GFP-histone 2B did not alter rates of chromosome mis-segregation (see Supplementary Table 1). In both cell lines, most binucleated cells (65–75%) arose from mononucleated cells that formed a bipolar spindle and initiated cytokinesis appropriately, but underwent furrow regression at a later time (see Supplementary Fig. 4). The delay between mitosis and furrow regression was long, averaging 13.6 h in HeLa cells, almost twice the average period between mitosis and furrow abscission (7.4 h) in cells that completed cytokinesis. In addition, 12–18% of binucleated cells arose from mononucleated cells that underwent a grossly abnormal mitosis. A small proportion of binucleated cells (3–6%) arose by fusion of two second-generation descendents of the same parent cell. Finally, 12% of binucleated cells arose through division of pre-existing binucleated cells. The genesis of binucleated cells is therefore complex, but the predominant pathway involves regression of the cleavage furrow in cells that undergo an otherwise normal-appearing mitosis.

Abnormal mitosis could potentially explain a high rate of chromo-

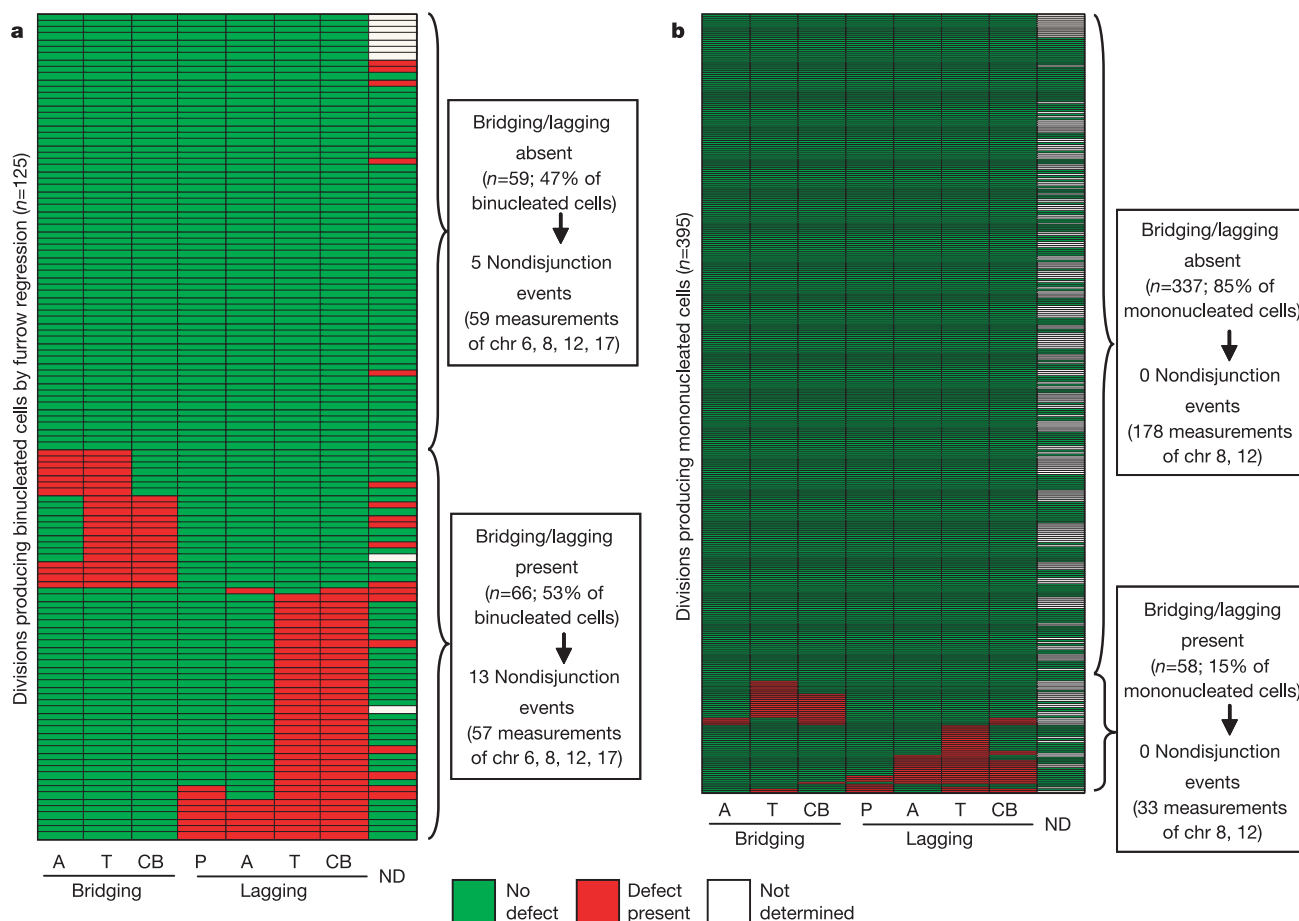


Figure 2 | Nondisjunction occurs with high frequency in cells that become binucleated by furrow regression, but not in cells that complete cytokinesis. HeLa cells expressing GFP-histone 2B were followed by time-lapse imaging and then fixed and analysed for chromosome segregation errors by FISH. **a**, Analysis of mononucleated cells that underwent furrow regression after bipolar mitosis to produce a binucleated cell. Each row represents analysis of a single cell division. Colour indicates whether the cell showed the presence (red) or absence (green) of chromosome bridging or

lagging, as determined by inspection of the movie, or nondisjunction, as measured by FISH. **b**, Analysis of mononucleated cells that completed cytokinesis after bipolar mitosis producing two mononucleated cells. The vertical scale of this plot is compressed due to the larger number of cells analysed. Abbreviations are as follows: A, anaphase; T, telophase; CB, interphase period following mitosis when cells are connected by a cytoplasmic bridge; P, prometaphase; ND, nondisjunction; and chr, chromosome.

some mis-segregation in binucleated cells, but only a small fraction of binucleated cells showed gross defects in mitosis. We therefore wanted to determine whether cells that divided with a normal bipolar spindle but became binucleated through furrow regression would also show an elevated rate of chromosome mis-segregation. We followed HeLa GFP-histone 2B cells by time-lapse microscopy, and then fixed the cells to analyse chromosome segregation by FISH. Using probes for chromosomes 6, 8, 12 and 17, we successfully measured chromosome segregation in 116 cells that became binucleated by furrow regression, and identified 18 cells that exhibited nondisjunction (Fig. 2a; see also Supplementary Table 2). In contrast, we detected no nondisjunction events in cells that completed cytokinesis to form two mononucleated cells (Fig. 2b; 211 measurements using probes for chromosomes 8 and 12). The average nondisjunction rate per chromosome in cells that became binucleated by furrow regression was 4.1% (see Supplementary Table 2), over ten times higher than that observed in the entire mitotic population (0.32%). If the rate of chromosome nondisjunction is similar for other chromosomes, then we predict that 94% of binucleated cells would show evidence of nondisjunction if segregation of all chromosomes was measured (see Supplementary Table 2).

These results indicate that regression of the cleavage furrow is tightly coupled to chromosome nondisjunction in human cells. One mechanism that could explain this finding is the presence of chromatin in the cleavage furrow, which has been reported to interfere with the completion of cytokinesis¹⁹. We therefore examined the time lapse movies from the previous experiments for the presence of chromosome bridging or lagging (CBL) during mitosis, and clustered the cells on this basis (Fig. 2). CBL was frequent in cells that underwent furrow regression (Fig. 2a; 53%), but was also present in a fraction of cells that completed cytokinesis (Fig. 2b; 15%). Although CBL is more frequent in cells that become binucleated, it appears neither necessary nor sufficient to explain binucleation. CBL was present in only half of all binucleated cells, and a significant number of cells completed cytokinesis despite the presence of CBL.

We believe that CBL is more frequent in binucleated cells because nondisjunction is more likely to occur in cells with CBL. For example, merotelic kinetochore attachment is a known cause of lagging chromosomes and nondisjunction in mammalian cells²⁰. However, merotelic attachment may not always result in nondisjunction if the lagging chromosome is ultimately moved to the correct pole, explaining why chromosome lagging may not be sufficient for binucleation.

If nondisjunction itself, rather than CBL, is coupled tightly to furrow regression, then the rate of nondisjunction in all mitotic cells should be similar to the rate of binucleation by furrow regression. We therefore determined the rate of binucleation by furrow regression in mononucleated cells that divided with a normal bipolar spindle (Fig. 3a, b). We found that the rates of binucleation, 1.9% per division for N/TERT-1 cells and 7.7% for HeLa cells, were very similar to the predicted rates of nondisjunction of all chromosomes during normal bipolar mitosis, that being 1.7% in N/TERT-1 cells and 7.3% in HeLa cells (see Supplementary Material for calculation). Thus, the rate of chromosome nondisjunction in cells with a normal bipolar mitosis is sufficient to account for the rate of binucleation by furrow regression, and a higher rate of chromosome nondisjunction in HeLa cells may explain the correspondingly higher rate of binucleation compared to N/TERT-1 cells.

To explore the potential consequences of binucleation, we compared the fates of mononucleated and binucleated cells by time-lapse imaging. During the period of imaging, we found that 86% of mononucleated N/TERT-1 cells divided, whereas only 40% of binucleated cells did so (Fig. 3c). In contrast, binucleated HeLa cells showed little delay in cell cycle progression compared to their mononucleated counterparts. We also determined whether dividing binucleated cells underwent mitosis with a bipolar or multipolar spindle (Fig. 3d). In binucleated N/TERT-1 cells that entered mitosis, 65% formed a bipolar spindle, with subsequent formation of two mononucleated daughter cells, and 35% formed multipolar spindles. In contrast, 94% of binucleated HeLa cells that entered mitosis

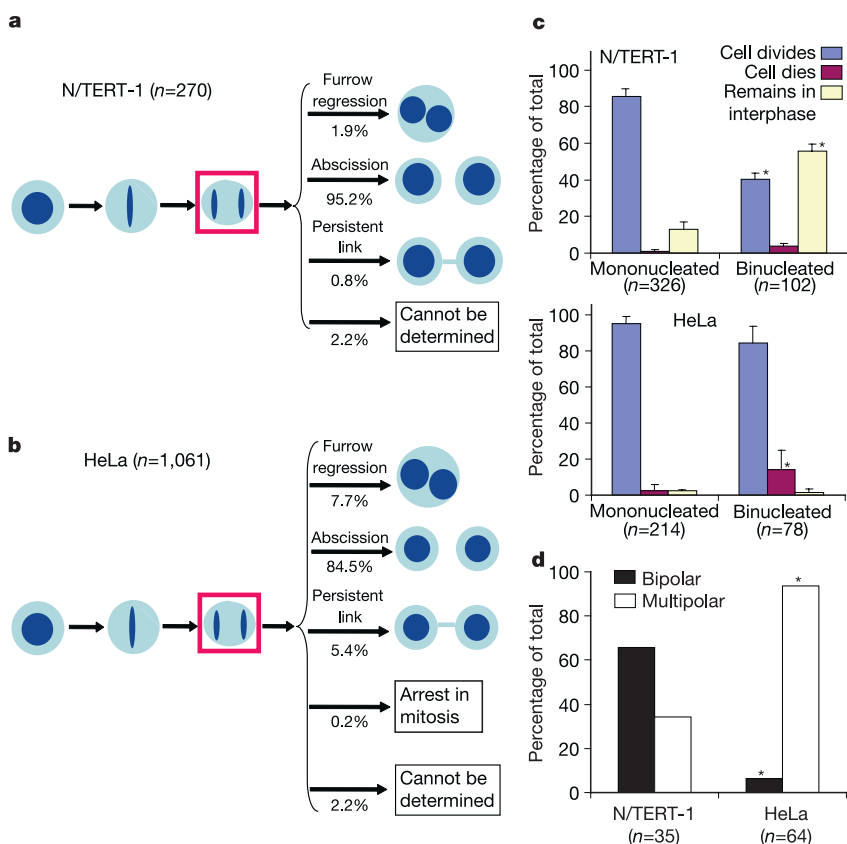


Figure 3 | Fates of mononucleated and binucleated N/TERT-1 and HeLa cells determined by long-term imaging. **a, b**, Mononucleated cells that underwent a bipolar mitosis followed by normal cytokinesis initiation were analysed by time-lapse imaging. The red box corresponds to the point at which cells were analysed for chromosome segregation by FISH in Fig. 1 and Supplementary Fig. 1. **c**, Analysis of division and apoptotic frequency of mononucleated and binucleated cells. Error bars are one standard deviation. **d**, Analysis of spindle polarity during mitotic division of binucleated cells. Asterisks indicate $P < 0.001$, $2 \times 2 \chi^2$ test, comparing frequencies in binucleated and mononucleated cells for both N/TERT-1 and HeLa cells (**c**) or comparing spindle polarity frequencies in binucleated N/TERT-1 and HeLa cells (**d**).

formed multipolar spindles, with only 6% forming a bipolar spindle. The products of the multipolar divisions included mononucleated, binucleated and multinucleated cells (see Supplementary Table 3). These results predict that mononucleated tetraploid cells should be present among N/TERT-1 cells but not HeLa cells. Indeed, measurements of the ploidy of interphase mononucleated N/TERT-1 cells by FISH indicated that 12% of these cells are tetraploid (see Supplementary Fig. 5). In contrast, only 0.2% of interphase mononucleated HeLa cells have a chromosome composition consistent with a 4N DNA content. These data indicate that mononucleated tetraploid cells can arise through a process involving chromosome nondisjunction and furrow regression, followed by a bipolar mitosis in the subsequent division.

Our data indicate that completion of cytokinesis requires accurate chromosome segregation in human cells (Fig. 4). Although the mechanistic basis for this association remains to be elucidated, our findings have significant implications for understanding how chromosome segregation errors are related to changes in ploidy. We propose that single nondisjunction events are tightly coupled to regression of the cleavage furrow, such that nondisjunction results in the formation of one binucleated cell rather than two mononucleated aneuploid cells. As suggested by our analysis of N/TERT-1 cells, cell division may be delayed in spontaneously arising binucleated cells, thus suppressing the accumulation of cells with mis-segregated chromosomes. As cells do not seem to possess a tetraploidy checkpoint^{21,22}, this delay may be a consequence of nondisjunction rather

than tetraploidization. If binucleated cells do divide, they appear to have the capacity to re-segregate the errant chromosome in the subsequent mitosis to produce two genetically equivalent tetraploid cells. However, binucleated cells also contain an additional centrosome, which may promote a multipolar mitosis, as was observed in HeLa cells. Thus, depending on the regulation of spindle and centrosome function in binucleated cells, coupling of nondisjunction with cytokinesis may also potentiate chromosome mis-segregation during subsequent rounds of division.

Although loss of spindle checkpoint proteins²³ or amplification of mitotic regulators such as Aurora A⁶ may be sufficient to induce chromosome segregation defects and cytokinesis failure, our findings suggest that such mutational events may not be required to initiate changes in ploidy in human cells. Instead, we propose that single stochastic errors in chromosome segregation may be sufficient to double cellular ploidy. The activation of oncogenes or loss of tumour suppressor genes might be necessary at a subsequent step, enabling binucleated cells to proliferate inappropriately, or to divide with a multipolar spindle to yield aneuploid progeny. Such a model might explain the correlation between loss of p53 and expansion of a tetraploid cell population in Barrett's oesophagus²⁴, a premalignant condition that precedes development of oesophageal cancer. Our findings are consistent with the notion that aneuploid cells may first proceed through a tetraploid state^{9,24–27}, and that spontaneous segregation errors may be sufficient to initiate a process that ultimately leads to aneuploidy²⁸. The tight coupling of nondisjunction with cytokinesis suggests that tetraploid cells may arise more frequently than previously thought, and may explain why single gains or losses of chromosomes appear to be relatively rare in cancers, whereas a pattern of genome duplication followed by chromosome loss seems common²⁹.

METHODS

Cell lines. N/TERT-1 cells were cultured according to published conditions¹⁴ and were found to have a normal male karyotype. PreEC cells (LH derivative line) were cultured as described¹⁶. HeLa cells (CCL2 from ATCC) were cultured in DMEM supplemented with 10% fetal calf serum and 1% non-essential amino acids (Invitrogen). HeLa cells had a modal chromosome number of 82, with 4 copies of chromosome 12 and 3 copies of chromosomes 6, 8 and 17. HeLa and N/TERT-1 cells expressing GFP–histone 2B were created as described in Supplementary Methods. Rates of chromosome mis-segregation and binucleation were similar between the parental cell lines and their GFP–histone 2B-expressing derivatives (see Supplementary Table 1).

Fluorescence *in situ* hybridization. Probes were labelled, hybridized and analysed by fluorescence microscopy as described³⁰ (see Supplementary Methods). For the data presented in Fig. 2 and Supplementary Table 2, coverslips were processed twice using two different sets of probes. Non-overlapping, intact cells were analysed for chromosome mis-segregation. For the data presented in Fig. 1 and Supplementary Fig. 1, mitotic and binucleated cells were analysed from the same regions of the coverslip. For mitotic cells, only anaphase or telophase cells with normal bipolar morphology were analysed. Interphase cells were verified for binuclear status using differential interference contrast microscopy. Binucleated cells were scored only if there was no overlap of the two nuclei or no hybridization signal in the overlap area if the two nuclei overlapped (see Supplementary Fig. 6). Nuclei were scored as having two or more copies of a specific chromosome if the signals of the same colour were of similar size and intensity and separated by a distance of more than half the diameter of the spot. To eliminate artefacts (for example, close, overlapping, missing or split signals; see Supplementary Fig. 6), only cells with an even total number of hybridization signals for every chromosome were scored. If the number of hybridization signals for a specific chromosome in each of the two nuclei of a cell was different, the cell was scored as having nondisjunction. Cells that underwent endomitosis were excluded from analysis based on their large nuclear size compared to neighbouring cells (see Supplementary Fig. 6).

Live cell imaging and analysis. HeLa or N/TERT-1 cells expressing GFP–histone 2B were grown on gridded coverglass bottom dishes (MatTek, Cat No. P35G-1.5-7-C-grid) for 26 h (HeLa, 2×10^5 cells per dish) or 3 d (N/TERT-1, 4×10^4 cells per dish) before imaging. Images were acquired automatically at multiple locations on the coverslip using a Nikon TE2000E inverted microscope fitted with a $\times 20$ Nikon Plan Fluor objective, a linearly-encoded stage (Proscan, Prior)

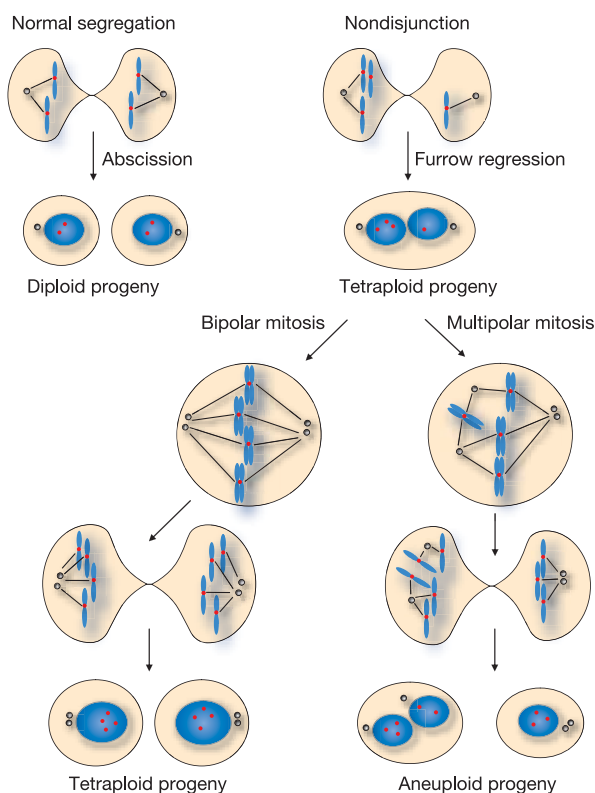


Figure 4 | Model summarizing the relationship of chromosome mis-segregation in the regulation of cytokinesis, and subsequent possible fates of resulting binucleated cells. DNA is blue and red spots represent centromeric FISH probes. Normal segregation (left) is associated with completion of cytokinesis producing two diploid mononucleated daughter cells. Chromosome nondisjunction (right) is associated with furrow regression, producing a binucleated tetraploid cell. If this cell divides, it can proceed through bipolar mitosis to produce two mononucleated tetraploid cells with equivalent genomes. However, if a multipolar spindle is formed, aneuploid progeny are likely to be produced due to high rates of chromosome mis-segregation resulting from multipolar mitosis.

and a Hamamatsu Orca-ER CCD camera. Fluorescence illumination used a mercury-arc lamp with two neutral density filters (for a total 32-fold reduction in intensity). The microscope was controlled using Simple PCI (Compix). The microscope was housed in a custom-designed 37 °C chamber with a secondary internal chamber that delivered humidified 5% CO₂. Fluorescence and differential interference contrast images were obtained every 15–22.5 min for a period of 28–88 h. Autofocusing was performed every 90 min using the fluorescence channel. Imaging conditions did not adversely affect cell proliferation or viability or increase the rate of binucleation compared to unimaged controls (data not shown).

For the experiment in Fig. 2, HeLa GFP–histone 2B cells were imaged for periods of 64, 68 and 88 h in three separate experiments, and then the coverslips were processed for FISH. Figure 2 illustrates the combined results of the three experiments; comparable results were obtained in each experiment. Cells were analysed if they entered mitosis at least 20 h before the end of imaging. Mononuclear cells that divided with a bipolar mitotic spindle were analysed for the presence of chromosome bridges or lagging chromosomes. We also documented whether cells completed cytokinesis or underwent regression of the cleavage furrow. The cells were scored as having chromosome bridges if the GFP–histone 2B positive nuclear material extended continuously between the two masses of chromosomes in anaphase or telophase, or between two daughter nuclei in the subsequent interphase (see Supplementary Fig. 7). Cells were considered to have lagging chromosomes if GFP–histone 2B-positive chromatin was observed in the midzone in anaphase or telophase, or in the cytoplasmic bridge in interphase, without evidence of chromatin linking the two daughter nuclei. Based on the presence or absence of chromosome lagging or bridging, the cells were then clustered hierarchically using Spotfire (UPGMA method using correlation as similarity measure). Results were displayed in Spotfire as a heat map.

For the experiments in Fig. 3a, b, cells expressing GFP–histone 2B were imaged for periods of 42–48 h (N/TERT-1 cells) and 28–72 h (HeLa cells). For Fig. 3c–e, cells were imaged for periods of 42–52 h (N/TERT-1 cells) or 28–72 h (HeLa cells). Standard deviations were derived from the results from 5 (binucleated N/TERT-1), 2 (mononucleated N/TERT-1), 4 (binucleated HeLa) or 2 (mononucleated HeLa) experiments.

Received 14 April; accepted 22 June 2005.

- Cahill, D. P. *et al.* Mutations of mitotic checkpoint genes in human cancers. *Nature* **392**, 300–303 (1998).
- Fodde, R. *et al.* Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nature Cell Biol.* **3**, 433–438 (2001).
- Hernando, E. *et al.* Rb inactivation promotes genomic instability by uncoupling cell cycle progression from mitotic control. *Nature* **430**, 797–802 (2004).
- Jallepalli, P. V. *et al.* Securin is required for chromosomal stability in human cells. *Cell* **105**, 445–457 (2001).
- Kaplan, K. B. *et al.* A role for the Adenomatous Polyposis Coli protein in chromosome segregation. *Nature Cell Biol.* **3**, 429–432 (2001).
- Meraldi, P., Honda, R. & Nigg, E. A. Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53^{-/-} cells. *EMBO J.* **21**, 483–492 (2002).
- Rajagopalan, H. *et al.* Inactivation of hCDC4 can cause chromosomal instability. *Nature* **428**, 77–81 (2004).
- Nigg, E. A. Centrosome aberrations: cause or consequence of cancer progression? *Nature Rev. Cancer* **2**, 815–825 (2002).
- Storchova, Z. & Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nature Rev. Mol. Cell Biol.* **5**, 45–54 (2004).
- Bharadwaj, R. & Yu, H. The spindle checkpoint, aneuploidy and cancer. *Oncogene* **23**, 2016–2027 (2004).
- Draviam, V. M., Xie, S. & Sorger, P. K. Chromosome segregation and genomic stability. *Curr. Opin. Genet. Dev.* **14**, 120–125 (2004).
- Carere, A. *et al.* Analysis of chromosome loss and non-disjunction in cytokinesis-blocked lymphocytes of 24 male subjects. *Mutagenesis* **14**, 491–496 (1999).
- Cimini, D., Tanzarella, C. & Degross, F. Differences in malsegregation rates obtained by scoring ana-telophases or binucleate cells. *Mutagenesis* **14**, 563–568 (1999).
- Dickson, M. A. *et al.* Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics. *Mol. Cell. Biol.* **20**, 1436–1447 (2000).
- Macville, M. *et al.* Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. *Cancer Res.* **59**, 141–150 (1999).
- Berger, R. *et al.* Androgen-induced differentiation and tumorigenicity of human prostate epithelial cells. *Cancer Res.* **64**, 8867–8875 (2004).
- Tanabe, K., Ikegami, Y., Ishida, R. & Andoh, T. Inhibition of topoisomerase II by antitumor agents bis(2,6-dioxopiperazine) derivatives. *Cancer Res.* **51**, 4903–4908 (1991).
- Kanda, T., Sullivan, K. F. & Wahl, G. M. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.* **8**, 377–385 (1998).
- Mullins, J. M. & Bieseke, J. J. Terminal phase of cytokinesis in D-98s cells. *J. Cell Biol.* **73**, 672–684 (1977).
- Cimini, D. *et al.* Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J. Cell Biol.* **153**, 517–527 (2001).
- Uetake, Y. & Sluder, G. Cell cycle progression after cleavage failure: mammalian somatic cells do not possess a “tetraploidy checkpoint”. *J. Cell Biol.* **165**, 609–615 (2004).
- Wong, C. & Stearns, T. Mammalian cells lack checkpoints for tetraploidy, aberrant centrosome number, and cytokinesis failure. *BMC Cell Biol.* **6**, 6 (2005).
- Kops, G. J., Foltz, D. R. & Cleveland, D. W. Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint. *Proc. Natl Acad. Sci. USA* **101**, 8699–8704 (2004).
- Galipeau, P. C. *et al.* 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett’s esophagus. *Proc. Natl Acad. Sci. USA* **93**, 7081–7084 (1996).
- Burholt, D. R. *et al.* Karyotypic evolution of a human undifferentiated large cell carcinoma of the lung in tissue culture. *Cancer Res.* **49**, 3355–3361 (1989).
- Kaneko, Y. & Knudson, A. G. Mechanism and relevance of ploidy in neuroblastoma. *Genes Chromosom. Cancer* **29**, 89–95 (2000).
- Shackney, S. E. *et al.* Model for the genetic evolution of human solid tumors. *Cancer Res.* **49**, 3344–3354 (1989).
- Duesberg, P. & Li, R. Multistep carcinogenesis: a chain reaction of aneuploidizations. *Cell Cycle* **2**, 202–210 (2003).
- Mitelman, F., Johansson, B. & Mertens, F. (eds) Mitelman database of chromosome alterations in cancer. (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>) (2005).
- Shi, Q. *et al.* Increased nondisjunction of chromosome 21 with age in human peripheral lymphocytes. *Mutat. Res.* **452**, 27–36 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Rheinwald for N/TERT-1 cells, W. Hahn for PrEC cells, J. Waters and the Nikon Imaging Center at Harvard Medical School for assistance and equipment, T. Mitchison for discussions, and D. Pellman, A. Amon, D. Moazed and P. Jackson for comments on the manuscript. This work was supported by the Harry C. McKenzie Family Foundation and the Harvard-Armenise Foundation. R.W.K. is a Damon Runyon-Walter Winchell Foundation Scholar.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.K. (randy_king@hms.harvard.edu).

Cytokinesis failure generating tetraploids promotes tumorigenesis in *p53*-null cells

Takeshi Fujiwara^{1*}, Madhavi Bandi^{1*}, Masayuki Nitta¹, Elena V. Ivanova², Roderick T. Bronson⁴
& David Pellman^{1,3}

A long-standing hypothesis on tumorigenesis is that cell division failure, generating genetically unstable tetraploid cells, facilitates the development of aneuploid malignancies^{1–3}. Here we test this idea by transiently blocking cytokinesis in *p53*-null (*p53*^{−/−}) mouse mammary epithelial cells (MMECs), enabling the isolation of diploid and tetraploid cultures. The tetraploid cells had an increase in the frequency of whole-chromosome mis-segregation and chromosomal rearrangements. Only the tetraploid cells were transformed *in vitro* after exposure to a carcinogen. Furthermore, in the absence of carcinogen, only the tetraploid cells gave rise to malignant mammary epithelial cancers when transplanted subcutaneously into nude mice. These tumours all contained numerous non-reciprocal translocations and an 8–30-fold amplification of a chromosomal region containing a cluster of matrix metalloproteinase (MMP) genes. MMP overexpression is linked to mammary tumours in humans and animal models⁴. Thus, tetraploidy enhances the frequency of chromosomal alterations and promotes tumour development in *p53*^{−/−} MMECs.

Many aneuploid tumour cells contain multiple centrosomes^{2,5–7}. The idea that an abnormal mitosis in a cell containing multiple centrosomes (a multipolar mitosis) would lead to genomic instability and cancer dates back almost 100 years to Theodor Boveri¹. Boveri proposed that multiple centrosomes could arise from either abnormal centrosome replication or a failure of cell cleavage leading to tetraploidy (so-called ‘double-value cells’). More recently, the notion that a tetraploid intermediate might have a role in oncogenesis has received support from studies of tissue culture cells^{8,9}, animal models¹⁰ and, most importantly, from human patients^{3,11,12}. For example, in patients with a premalignant condition called Barrett’s oesophagus, tetraploid cells are detected before gross aneuploidy, but close to the time at which the tumour suppressor gene *p53* is lost¹¹. Despite these intriguing correlations, in the systems where tetraploidy has been studied so far, it has not been possible to directly demonstrate the significance of tetraploid cells in tumorigenesis.

To determine if the formation of tetraploid epithelial cells can promote tumorigenesis, we developed conditions for the culture of genetically matched diploid and tetraploid MMECs. We focused on cells lacking *p53* because (1) there is an epidemiological link between *p53* loss and tetraploidy in Barrett’s oesophagus¹¹, (2) *p53* loss facilitates the spontaneous formation of tetraploid cells in a variety of mouse and human cell types in culture¹³, and (3) *p53* loss is often required for mutations affecting genomic stability to manifest as tumours^{14,15}.

Cytokinesis was transiently blocked in primary MMECs isolated from wild-type (*p53*^{+/+}) or *p53*^{−/−} animals by treatment with 2 μ M

dihydrocytochalasin B (DCB), a compound that prevents remodeling of the actin cytoskeleton. After 20 h of DCB treatment, 60% of the *p53*^{−/−} MMECs and 52% of the *p53*^{+/+} MMECs were binucleate, indicating a failure of cytokinesis¹⁶ (Fig. 1a and data not shown). Both populations of cells proceeded through the cell cycle and entered mitosis, as evidenced by the rapid appearance of an 8C (G2/M phase tetraploid) peak on flow cytometry and the labelling of cells to visualize mitotic figures (Fig. 1a, Supplementary Figs 1 and 2). Therefore, primary MMECs may not have a checkpoint that triggers a long-term G1 arrest in response to tetraploidy⁸, consistent with recent findings in other cell types^{17,18}.

Next, diploid and tetraploid MMECs were isolated by fluorescence-activated cell sorting (FACS), separating the 2C (diploid cells in G1 phase) from the 8C (tetraploid cells in G2/M phase) peak (Fig. 1a). Diploid and tetraploid cultures of *p53*^{−/−} cells were readily propagated (Supplementary Fig. 1c), grew at equivalent rates (doubling times of 25 h and 26 h, respectively), and contained similar percentages of 5-bromodeoxyuridine (BrdU)-labelled cells (Supplementary Fig. 1d). Notably, although aneuploidy developed with the subsequent propagation of both *p53*^{−/−} diploids and tetraploids, chromosome spreads, spectral karyotyping (SKY) and comparative genomic hybridization (array-CGH) demonstrated that both cultures predominantly contained the expected numbers of normal chromosomes immediately after FACS (at metaphase, diploids have 40 chromosomes and tetraploids have 80 chromosomes; Fig. 1b, c, Supplementary Fig. 3). In contrast, although the *p53*^{+/+} diploids could be propagated, the *p53*^{+/+} tetraploids failed to propagate in culture (Supplementary Fig. 2). The ability to isolate and propagate diploid and tetraploid *p53*^{−/−} MMECs enabled us to characterize their genetic stability and susceptibility to transformation.

Tetraploid cells are known to produce multipolar mitoses that result in whole-chromosome mis-segregation (whole-chromosome aneuploidy)^{1,8}. However, it is not known if tetraploidy can facilitate gross chromosomal rearrangements (GCRs). Consistent with previous studies of primary murine *p53*^{−/−} cells¹³, cultures derived from diploid and tetraploid *p53*^{−/−} MMECs developed whole-chromosome aneuploidy. However, at each time point after initial culture, the tetraploid-derived *p53*^{−/−} cells showed more whole-chromosome aneuploidy than the diploid-derived cells (Fig. 1b). SKY showed that the tetraploid-derived cultures of *p53*^{−/−} cells also contained more cells with GCRs than the diploid-derived cultures ($P < 0.02$, Student’s *t*-test; Fig. 1c). These experiments are likely to underestimate the difference in the amount of aneuploidy between the diploid and tetraploid cultures because the diploid cultures spontaneously become tetraploid during passage in culture (Supplementary Fig. 1e and data not shown). Thus, tetraploidy increases

¹Department of Pediatric Oncology and ²Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, ³Pediatric Hematology/Oncology, Children’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁴Department of Biomedical Sciences, Tufts University Veterinary School, North Grafton, Massachusetts 01536, USA.

*These authors contributed equally to this work.

the proportion of cells with both whole-chromosome aneuploidy and GCRs in cultures of $p53^{-/-}$ MMECs.

Next, the impact of tetraploidy on *in vitro* transformation was investigated by assaying the growth of the cells in soft agar (Fig. 2a, schematic 1). Neither the diploid- nor the tetraploid-derived $p53^{-/-}$ MMECs showed anchorage-independent growth after treatment with the vehicle controls, indicating that formation of tetraploids alone is not sufficient to induce transformation (Fig. 2b and data not shown). However, in an initiation/promotion transformation assay¹⁹, where cells were exposed to the mutagen 7,12-dimethylbenz[*a*]anthracene (DMBA) followed by exposure to the tumour promoter 12-*O*-tetradecanoylphorbol-13-acetate (TPA), only the tetraploid cells showed anchorage-independent growth (Fig. 2b). This finding was observed in three independent experiments using three different donor animals. Unlike control MDA-MB-435 cells, which maintain discrete colony morphology, the transformed tetraploid-derived cells migrated through the agar and rapidly grew to confluence (Fig. 2b). The tetraploid-derived cells growing in soft agar produced tumours within two weeks when injected subcutaneously into nude mice (in 8 out of 8 animals; data not shown), thus demonstrating that they are transformed. SKY revealed extensive GCRs in the transformed tetraploid-derived cells. Out of 21 cells examined, 19 cells had non-reciprocal translocations, with a median of two translocations per cell (Supplementary Fig. 6).

The *in vitro* transformation of the tetraploid cells is unlikely to have resulted from the transient exposure to DCB, because the diploids and tetraploids were derived from the same culture and were exposed to DCB for the same period of time. Furthermore, DCB treatment induces neither DNA damage (Supplementary Fig. 4) nor whole-chromosome mis-segregation (ref. 20, and R. King, unpublished results). FACS also did not influence these results because our FACS procedure did not induce detectable DNA damage (Supplementary Fig. 5). Moreover, treatment with DCB alone, with no

FACS, was sufficient for transformation after exposure to DMBA and TPA (Fig. 2a, schematic 2 and Fig. 2c).

Next, we asked whether tetraploidy could promote tumorigenesis *in vivo* without carcinogen exposure. After DCB treatment, diploid and tetraploid $p53^{-/-}$ MMECs were isolated by FACS and 5×10^5 cells were injected contralaterally into nude mice without further *in vitro* culture—conditions where the cell populations predominantly contained the expected numbers of normal chromosomes (Fig. 1b, c, Supplementary Fig. 3). Out of 39 animals injected, ten developed tumours at the site where tetraploid $p53^{-/-}$ MMECs had been introduced, with a latency of 9–12 weeks (Fig. 3a). In contrast, no tumours were observed at the site where diploid cells were injected ($P < 0.01$, Student's *t*-test). The ten tumours arose from five independent experiments using cells from different donor animals. Histopathological analysis, including immunohistochemistry, showed that all nine of the tumours analysed were malignant myoepitheliomas (Fig. 3b–d and Supplementary Fig. 7). Consistent with their derivation from tetraploid cells, immunofluorescence labelling for pericentrin demonstrated that at least 65% of the tumour cells contain extra centrosomes (Fig. 3e).

The tetraploid-derived tumours manifested numerical and structural chromosome aberrations. For the two tumours analysed, the median metaphase chromosome numbers were 89 and 74 (Supplementary Fig. 8a). Both tumours contained abundant non-reciprocal translocations (NRTs), dicentric chromosomes and double minute chromosomes, with a median of two chromosomal rearrangements per cell (Fig. 4a and Supplementary Fig. 8b).

Array-CGH provided further insight into the genetic alterations in these tumours. Each of the nine tumours examined had an 8–30-fold amplification of the 7.16–7.98-Mb region of chromosome 9, which contains a cluster of 10 MMP genes (Fig. 4b, c, Supplementary Figs 9 and 10). Notably, the amplicon we identified contains *Mmp1*, *Mmp3*, *Mmp7* and *Mmp13*—genes that are all overexpressed in human breast tumours⁴. Furthermore, transgenic animals overexpressing

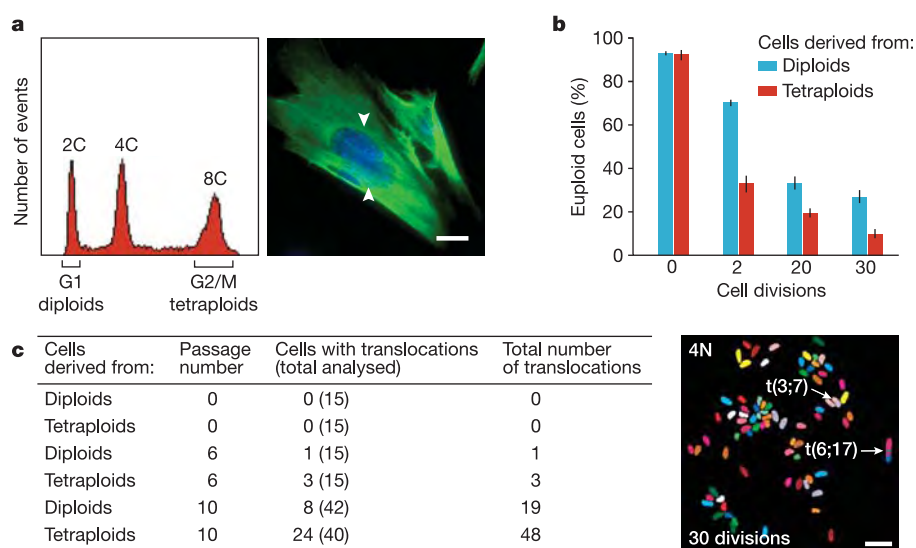


Figure 1 | Tetraploid $p53^{-/-}$ MMECs are viable, but develop numerical and structural chromosomal abnormalities after passage. **a**, Isolation of tetraploid $p53^{-/-}$ MMECs. Left, FACS profiles after 20 h of DCB treatment. Right, immunofluorescence image showing cytokeratin expression (green) and DAPI-labelling of DNA (blue). A binucleate cell (arrowhead indicates nuclei) after DCB treatment is shown. Scale bar, 10 μ m. **b**, Tetraploid MMECs lose the euploid chromosome content to a greater extent than diploid MMECs with passage (the bars indicate the percentage of euploid cells in the culture). The number of cell divisions is based on the doubling rate; cell division 20 corresponds to passage 6 and cell division 30

corresponds to passage 10. The results are the average of three independent experiments where at least 50 metaphase spreads were analysed for each sample. Error bars indicate standard deviations. **c**, Left, a table showing increased NRTs in both the diploid- and tetraploid-derived $p53^{-/-}$ MMECs from passage 0 to passage 10 (3-fold increase, $P < 0.02$). Passage 0 corresponds to analysis done 48 h after seeding cells onto plates after cell sorting. Right, representative SKY showing NRTs in a late-passage (passage 10) tetraploid-derived MMEC. Note, these translocations are all nonrecurrent. Scale bar, 10 μ m.

Mmp3 or *Mmp7* in the mammary gland develop either epithelial hyperplasia or tumours⁴. Three tumours also contained a 4–8-fold amplification of the 16.58–18.18-Mb region of chromosome 6 containing the proto-oncogene *c-Met* (Supplementary Fig. 9), which is also overexpressed in many human breast cancers²¹. Thus, the tumours derived from tetraploid *p53*^{-/-} MMECs contained genetic alterations suggesting similarities to human breast cancers.

Our study demonstrates that a transient block in cytokinesis, leading to the formation of genetically unstable tetraploid cells, promotes tumorigenesis in *p53*^{-/-} mouse mammary epithelial cells. Variations of this idea have been discussed since the original versions of the hypothesis advanced by Theodor Boveri in 1914. Our results are consistent with, and provide functional support for, reports identifying roles for tumour suppressors or oncogenes in cytokinesis^{22,23}.

The frequency at which tetraploidy occurs in normal tissues is unknown but merits further investigation. However, there are reasons to conclude that it may occur at a significant rate, especially in cells that concurrently lose *p53*. Tetraploidy can be generated from a wide variety of errors in cell division^{2,3}. Even a subtle cell division defect, chromosome nondisjunction, has been recently associated with a high frequency of tetraploidy during *in vitro* culture²⁰. Furthermore, there may be a selective advantage for polyploid cells in certain conditions; for example, during times of metabolic stress³.

Understanding the mechanism by which tetraploidy leads to aneuploidy, particularly chromosomal breaks, will be an important topic of future work. Tetraploidy may increase the proportion of cells with chromosomal breaks directly, by increasing the rate of chromosome breaks, or indirectly, by enhancing the fitness of cells with chromosomal aberrations (because of the additional copies of complementing normal chromosomes). The aneuploidy arising from tetraploid cells is probably due, at least in part, to multiple centrosomes and multipolar mitoses. A multipolar mitosis could induce chromosome breaks if sister chromatid cohesion or catenation is not fully resolved. A multipolar mitosis could also lead to chromosome breaks indirectly: whole chromosome losses or gains could lead to gene expression imbalances, generating a defective

S phase, and thus producing DNA breakage. Finally, tetraploid cells may develop aneuploidy by mechanisms that do not involve extra centrosomes. In budding yeast, polyploidy (in the absence of extra centrosomes) leads to significant genetic instability (refs 3, 24, 25, and Z. Storchova and D.P., unpublished results).

There are interesting parallels between tetraploidy and telomere attrition that could indicate similar roles in tumorigenesis. In the situation where *p53* is lost, telomere attrition leads to epithelial cancers with significant aneuploidy¹⁴. However, in other cell types or genetic backgrounds, telomere attrition results in cell death and inhibits tumorigenesis²⁶. Similarly, although tetraploidy facilitated tumorigenesis in *p53*^{-/-} MMECs, tetraploidy inhibits the formation of teratomas by *p53*^{+/+} mouse embryonic stem cells (T.F., M.B. and

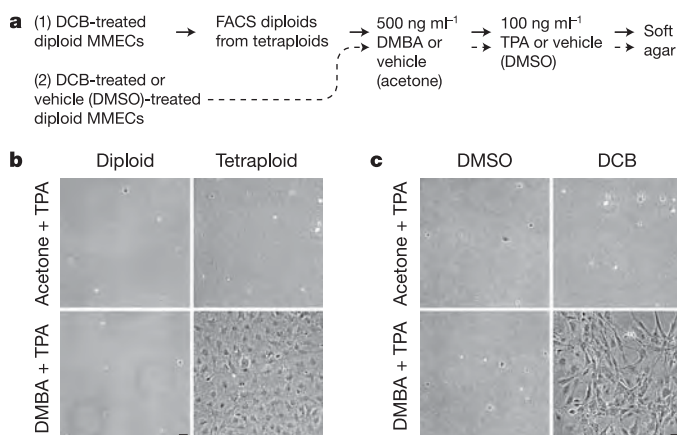


Figure 2 | Transformation of tetraploid *p53*^{-/-} MMECs after carcinogen treatment. **a**, Design of the two related *in vitro* experiments. **b**, **c**, Phase contrast images (**b** shows schematic 1 in **a**; **c** shows schematic 2 in **a**) of the indicated cell populations in soft agar. Note that the response to DMBA was dose-dependent, because there was no transformation in any of these conditions after treatment with 100 ng ml⁻¹ of DMBA (data not shown) in contrast with the 500 ng ml⁻¹ used in the experiments presented. Control experiments with the vehicle DMSO rather than TPA are not shown. DMBA is required for the transformation of tetraploid *p53*^{-/-} MMECs, but TPA is not (data not shown). Scale bars, 20 μ m.

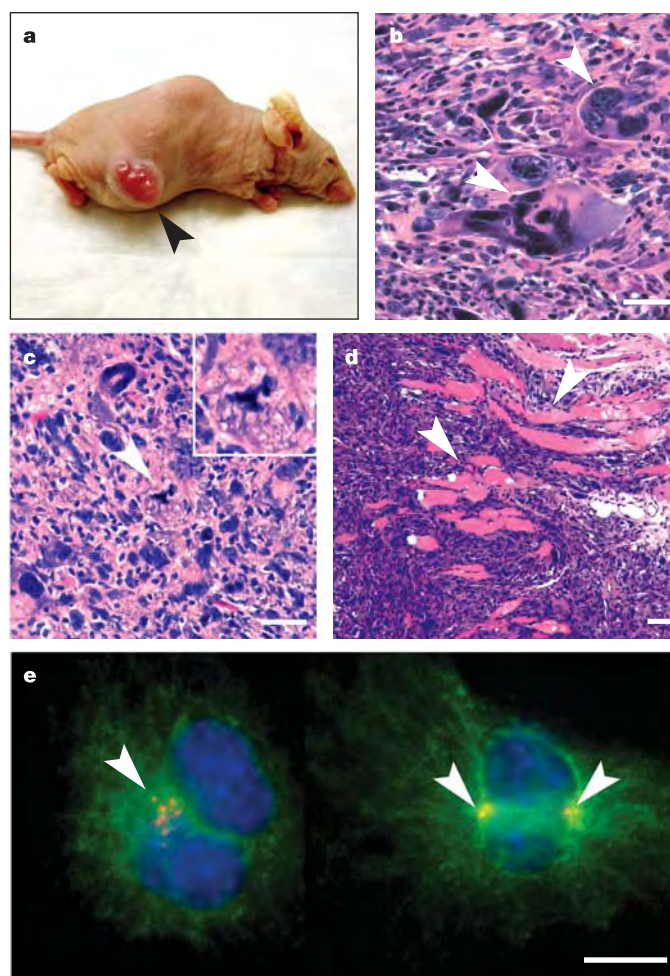


Figure 3 | Tetraploid *p53*^{-/-} MMECs spontaneously generate malignant myoepitheliomas after subcutaneous transplantation into nude mice. **a**, Tumour in a nude mouse at the site of injection of 5×10^5 tetraploid *p53*^{-/-} MMECs. **b–d**, Histology of the tumours derived from tetraploid MMECs (haematoxylin and eosin stain). The cell morphology is consistent with a malignant spindle-cell carcinoma. See Supplementary Fig. 7 for immunohistochemistry showing the labelling of tumours for smooth muscle actin and keratin 14, which suggests that they are myoepitheliomas. There is marked nuclear atypia, with frequent multinucleated giant cells (**b**, arrowhead), multipolar mitoses (**c**, arrowhead and inset) and tumour invasion of skeletal muscle (**d**, arrowhead). Scale bars, 100 μ m. **e**, Multiple centrosomes in a tetraploid-derived malignant myoepithelioma. Cells were labelled with pericentrin antibody (red, arrowhead) and α -tubulin antibody (green). Tumours were dissected, dispersed with collagenase A, and cultured for 5 days. At least 700 cells were counted in each tumour (nine tumours were examined). Scale bar, 10 μ m.

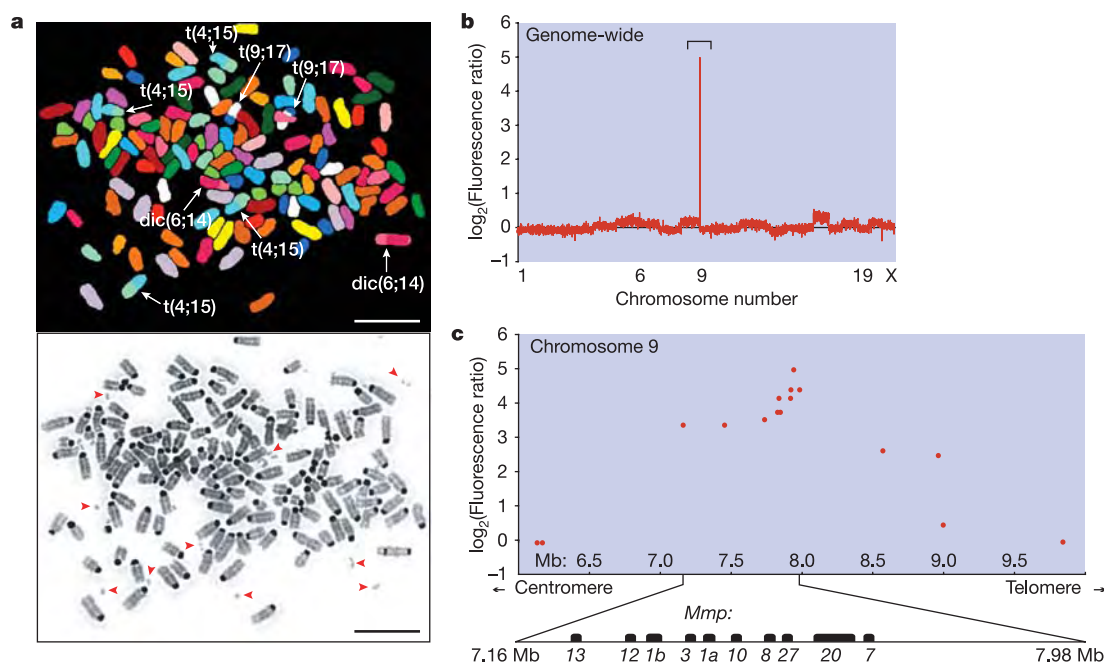


Figure 4 | Gross chromosomal rearrangements in tumours from tetraploid $p53^{-/-}$ MMECs. **a**, Representative SKY data from one tumour, showing NRTs (arrows, top), dicentric chromosomes (dic, top) and double minute chromosomes (red arrowheads in DAPI-stained sample, bottom). Scale bars, 10 μ m. **b**, Genome-wide array-CGH analysis reveals amplification of a 0.9-Mb region on chromosome 9 containing the tandem cluster of ten *Mmp* genes. The data shown is from one tumour; see the Supplementary

Information for data from the other eight tumours examined (Supplementary Fig. 9). The list of genes in the amplicon are shown in Supplementary Fig. 10. The y axis shows the $\log_2$ of the fluorescence intensity ratio, [Cy3 (tumour)/Cy5 (normal $p53^{-/-}$ diploid)]. A physical map of chromosome position is shown on the x axis. **c**, High resolution view of the amplicon from the same tumour illustrating the position of the *Mmp* genes on chromosome 9.

D.P., unpublished results). Indeed, some cell types, and even whole organisms, tolerate tetraploidy remarkably well^{3,27}. Therefore, like telomere attrition, the biological consequences of tetraploidy may be dependent on cell type and genetic context.

METHODS

Isolation and culture of tetraploid mouse mammary epithelial cells. Mouse mammary glands were isolated from 8–9-week-old female $p53^{+/+}$ and $p53^{-/-}$ littermates. Procedures for the isolation of mammary organoids and digestion with collagenase A to obtain MMECs were as previously described²⁸. MMECs were cultured on 0.1% (w/v) gelatin-coated plates (Speciality Media) with MECL growth medium, which consists of DMEM/F-12 (Gibco/Invitrogen) supplemented with 2% fetal bovine serum (Gibco), 10 μ g ml⁻¹ insulin (Sigma) and 5 ng ml⁻¹ epidermal growth factor (Sigma). MMECs were cultured for 5–6 days at 37 °C in 5% CO₂. Cells were then treated with 2 μ M DCB (Sigma) for 18–20 h. After DCB treatment, the cells were washed and incubated with growth medium for 1 h. Hoechst 33342 (Molecular Probes) was added to the culture medium (2 μ g ml⁻¹) 30 min before harvesting the cells. Tetraploid and diploid MMECs were FACS-sorted according to the DNA content, and seeded on 6- or 12-well culture dishes, or 4-well culture slides, coated with 0.1% (w/v) gelatin. The cells were subsequently passaged every three days.

Cytogenetic analyses. For conventional karyotyping, >50 metaphase spreads were analysed for each sample. SKY and array-CGH were performed as described^{29,30}. At least 15 metaphase spreads were analysed for each sample.

In vitro transformation assay. The protocol was adapted from previously described methods¹⁹. Sorted MMECs were seeded onto dishes coated with 0.1% (w/v) gelatin. After 12–18 h, the cells were treated with either acetone as vehicle or different concentrations of DMBA (Sigma) for 3 days. On day four, the cultures were placed in medium containing either dimethylsulphoxide (DMSO) as vehicle or 100 ng ml⁻¹ of TPA (Sigma) for an additional ten days. Equivalent numbers of diploid and tetraploid cells survived in culture after DMBA and TPA treatment. Cells (3,000) were mixed with 0.3% soft agar and layered on top of 0.6% bottom agar in 12-well dishes. In the experiment where treatment with either DMSO as vehicle or 2 μ M DCB was compared without FACS sorting, 6,000 cells were mixed with 0.3% soft agar.

In vivo tumorigenicity assay. Diploid and tetraploid $p53^{-/-}$ MMEC cultures were suspended in PBS at 5×10^5 cells per 150 μ l, and injected subcutaneously

into the contralateral sides of 39 NCR-*Foxn1*^{nu/nu} 8-week-old female mice (Taconic). All recipient animals were palpated weekly for tumour formation and euthanized when tumour masses were detected at 2 cm in diameter, in accordance with institutional procedures for animal safety and protection. The tumours were either sectioned for histopathological analysis or dissociated into single cells by collagenase A treatment as described above.

Histology and immunohistochemistry. Fixation and staining of histological sections with haematoxylin and eosin was performed according to standard protocols. For immunohistochemistry, tissue sections were labelled with the following antibodies: a monoclonal anti-pan-cytokeratin antibody (C25562, Sigma), a polyclonal anti-keratin-14 antibody (AF64, Covance), and a polyclonal anti-smooth-muscle-actin antibody (ab15267, Abcam).

Immunofluorescence and FACS. Samples for immunofluorescence were fixed with 4% PFA/PBS (pH 7.2–7.4) and permeabilized with 0.2% Triton-X100/PBS for 2 min. The following antibodies were used for immunofluorescence: a monoclonal anti-pan-cytokeratin antibody, a monoclonal anti- α -tubulin antibody (T9026, Sigma), and a polyclonal anti-pericentrin antibody. Images were captured with an Axioplan 2 microscope (Zeiss) equipped with an ORCA-ER digital camera (Hamamatsu). BrdU labelling was performed as described¹⁷. To prepare cells for FACS analysis, cultured MMECs were collected with 0.25% (w/v) Trypsin/EDTA (Gibco/Invitrogen). FACS data analysis was done using BD CellQuest Pro software (Becton Dickinson). The FACS conditions were set to minimize cell damage by the laser. A multiple-wavelength ultraviolet laser with mixed wavelengths of 351 nm and 361 nm (Coherent) was set at 40 mW, and the length of time of exposure was restricted to 2 μ s.

Received 9 May; accepted 5 September 2005.

1. Boveri, T. *The Origin of Malignant Tumors* (Williams and Wilkins, Baltimore, 1929).
2. Nigg, E. A. Centrosome aberrations: cause or consequence of cancer progression? *Nature Rev. Cancer* **2**, 815–825 (2002).
3. Storchova, Z. & Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nature Rev. Mol. Cell Biol.* **5**, 45–54 (2004).
4. Egeblad, M. & Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nature Rev. Cancer* **2**, 161–174 (2002).
5. Lingle, W. L. et al. Centrosome amplification drives chromosomal instability in breast tumour development. *Proc. Natl Acad. Sci. USA* **99**, 1978–1983 (2002).

6. Pihan, G. A., Wallace, J., Zhou, Y. & Doxsey, S. J. Centrosome abnormalities and chromosome instability occur together in pre-invasive carcinomas. *Cancer Res.* **63**, 1398–1404 (2003).
7. Jallepalli, P. V. & Lengauer, C. N. Chromosome segregation and cancer: cutting through the mystery. *Nature Rev. Cancer* **1**, 109–117 (2001).
8. Andreassen, P. R., Lohez, O. D., Lacroix, F. B. & Margolis, R. L. N. Tetraploid state induces p53-dependent arrest of nontransformed mammalian cells in G1. *Mol. Biol. Cell* **12**, 1315–1328 (2001).
9. Shackney, S. E. *et al.* Model for the genetic evolution of human solid tumors. *Cancer Res.* **49**, 3344–3354 (1989).
10. Levine, D. S., Sanchez, C. A., Rabinovitch, P. S. & Reid, B. J. Formation of the tetraploid intermediate is associated with the development of cells with more than four centrosomes in the elastase-simian virus 40 tumour antigen transgenic mouse model of pancreatic cancer. *Proc. Natl Acad. Sci. USA* **88**, 6427–6431 (1991).
11. Galipeau, P. C. *et al.* 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. *Proc. Natl Acad. Sci. USA* **93**, 7081–7084 (1996).
12. Olaharski, A. J. *et al.* Tetraploidy and chromosomal instability are early events during cervical carcinogenesis. *Carcinogenesis* published online, 25 August 2005 (doi:10.1093/carcin/bgi218).
13. Livingstone, L. R. *et al.* Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell* **70**, 923–935 (1992).
14. Artandi, S. E. *et al.* Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* **406**, 641–645 (2000).
15. Bassing, C. H. *et al.* Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114**, 359–370 (2003).
16. Carter, S. B. Effects of cytochalasins on mammalian cells. *Nature* **213**, 261–264 (1967).
17. Uetake, Y. & Sluder, G. Cell cycle progression after cleavage failure: mammalian somatic cells do not possess a "tetraploidy checkpoint". *J. Cell Biol.* **165**, 609–615 (2004).
18. Wong, C. & Stearns, T. Mammalian cells lack checkpoints for tetraploidy, aberrant centrosome number, and cytokinesis failure. *BMC Cell Biol.* **6**, 6 (2005).
19. Boutwell, R. K. The function and mechanism of promoters of carcinogenesis. *Crit. Rev. Toxicol.* **2**, 419–443 (1974).
20. Shi, Q. & King, R. W. Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. *Nature* doi:10.1038/nature03958 (this issue).
21. Birchmeier, C., Birchmeier, W., Gherardi, E. & Vande Woude, G. F. MET, metastasis, motility and more. *Nature Rev. Mol. Cell Biol.* **4**, 915–925 (2003).
22. Daniels, M. J., Wang, Y., Lee, M. & Venkitaraman, A. R. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science* **306**, 876–879 (2004).
23. Yang, X. *et al.* LATS1 tumour suppressor affects cytokinesis by inhibiting LIMK1. *Nature Cell Biol.* **6**, 609–617 (2004).
24. Mayer, V. W. & Aguilera, A. High levels of chromosome instability in polyploids of *Saccharomyces cerevisiae*. *Mutat. Res.* **231**, 177–186 (1990).
25. Lin, H. *et al.* Polyploids require Bik1 for kinetochore-microtubule attachment. *J. Cell Biol.* **155**, 1173–1184 (2001).
26. Feldser, D. M., Hackett, J. A. & Greider, C. W. Telomere dysfunction and the initiation of genome instability. *Nature Rev. Cancer* **3**, 623–627 (2003).
27. Stukenberg, P. T. Triggering p53 after cytokinesis failure. *J. Cell Biol.* **165**, 607–608 (2004).
28. Murphy, K. L., Dennis, A. P. & Rosen, J. M. A gain of function p53 mutant promotes both genomic instability and cell survival in a novel p53-null mammary epithelial cell model. *FASEB J.* **14**, 2291–2302 (2000).
29. Brennan, C. *et al.* High-resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Res.* **64**, 4744–4748 (2004).
30. David, G., Turner, G. M., Yao, Y., Protopopov, A. & DePinho, R. A. mSin3-associated protein, mSds3, is essential for pericentric heterochromatin formation and chromosome segregation in mammalian cells. *Genes Dev.* **17**, 2396–2405 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to S. Artandi, L. Chin, R. DePinho, C. M. Kuperwasser, M. E. McLaughlin, K. Polyak, A. Protopopov, J. Ruderman, J. Sage, P. Sicinski, T. Stearns and C. Wong for discussions; K. Polyak for showing us procedures with MMECs; A. D'Andrea, R. DePinho, M. Ewen, R. King, M. E. McLaughlin, K. Polyak and Z. Storchova for comments on the manuscript; J. Dunn for assistance with the figures; and S. Doxsey for the anti-pericentriolar antibody. T.F. was a Uehara Memorial Foundation research fellow. D.P. was supported by an NIH grant and a scholar award from the Leukemia and Lymphoma Society of America.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.P. (david_pellman@dfci.harvard.edu).

LETTERS

Calcium triggers exit from meiosis II by targeting the APC/C inhibitor XErp1 for degradation

Nadine R. Rau^{1*}, Andreas Schmidt^{1*}, Jenny Bormann¹, Erich A. Nigg² & Thomas U. Mayer¹

Vertebrate eggs awaiting fertilization are arrested at metaphase of meiosis II by a biochemical activity termed cytostatic factor (CSF)^{1,2}. This activity inhibits the anaphase-promoting complex/cyclosome (APC/C), a ubiquitin ligase that triggers anaphase onset and mitotic/meiotic exit by targeting securin and M-phase cyclins for destruction^{3,4,5}. On fertilization a transient rise in free intracellular calcium⁶ causes release from CSF arrest and thus APC/C activation. Although it has previously been shown that calcium induces the release of APC/C from CSF inhibition through calmodulin-dependent protein kinase II (CaMKII)^{7,8}, the relevant substrates of this kinase have not been identified. Recently, we characterized XErp1 (Emi2), an inhibitor of the APC/C and key component of CSF activity in *Xenopus* egg extract⁹. Here we show that calcium-activated CaMKII triggers exit from meiosis II by sensitizing the APC/C inhibitor XErp1 for polo-like kinase 1 (Plx1)-dependent degradation. Phosphorylation of XErp1 by CaMKII leads to the recruitment of Plx1 that in turn triggers the destruction of XErp1 by phosphorylating a site known to serve as a phosphorylation-dependent degradation signal. These results provide a molecular explanation for how the fertilization-induced calcium increase triggers exit from meiosis II.

Although it is well established that CaMKII is the essential target of the calcium signal^{7,8} on fertilization, the relevant substrates of this kinase have not been identified, and the mechanism(s) leading to APC/C activation have long remained obscure. Recently, we have identified XErp1, a novel component of CSF activity that is both necessary and sufficient to keep the APC/C inactive in CSF-arrested *Xenopus* egg extracts⁹ (CSF extracts). In response to calcium, XErp1 is rapidly degraded via a Plx1-dependent mechanism, leading to CSF release and APC/C activation. We have further shown that XErp1 destruction depends on phosphorylation of two critical serine residues within a motif (DSGX₃S) known to serve as a 'phospho-degron' for the ubiquitin ligase complex Skp1-Cullin-F-box^{β-TRCP} (SCF^{β-TRCP}; ref. 10). Although the identification of XErp1 as a critical substrate of Plx1 provided an attractive explanation for the essential function of Plx1 in APC/C activation¹¹, the role of the calcium signal remained unclear. In particular, it remained to be explained why XErp1 is not targeted for degradation in CSF extracts, despite the presence of active Plx1 and SCF^{β-TRCP}. Here we have explored the hypothesis that a calcium-dependent mechanism could sensitize XErp1 for phosphorylation by Plx1.

We first asked whether XErp1 could be a substrate of CaMKII.

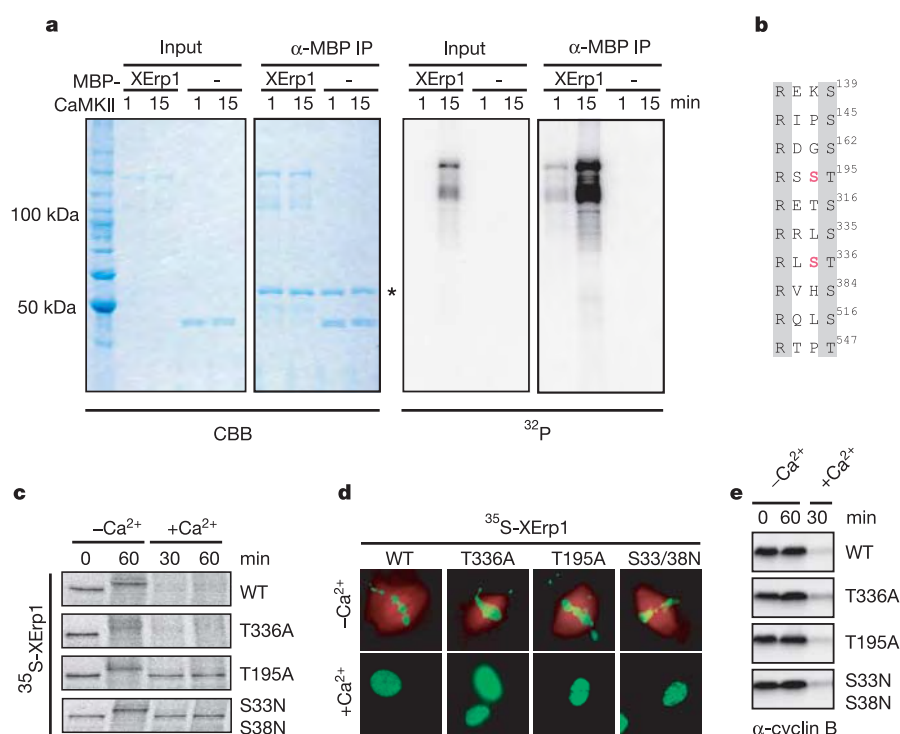


Figure 1 | XErp1 is a substrate of CaMKII in vitro.

a, MBP-tagged XErp1^{WT} or MBP alone were subjected to CaMKII phosphorylation reactions for the indicated times and immunopurified using anti-MBP antibodies (asterisk marks heavy chains). CBB indicates Coomassie Brilliant Blue. The incorporation of ³²P was analysed by autoradiography. **b**, Amino acid sequence of all putative CaMKII phosphorylation sites (RXXS/T) present in XErp1. **c–e**, Both DSGX₃S³⁸ and RXST¹⁹⁵ motifs are essential for the degradation of XErp1 on calcium-stimulation. ³⁵S-labelled IVT XErp1 proteins were incubated in CSF extract in the presence or absence of calcium. At the indicated time points samples were withdrawn for analysis by autoradiography (**c**), microscopic examination of chromatin and spindle structures (**d**) and immunoblotting for cyclin B (**e**).

¹Chemical Biology, Independent Research Group, ²Department of Cell Biology, Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany.

*These authors contributed equally to this work.

Autoradiographic analyses revealed that CaMKII efficiently phosphorylated purified maltose-binding protein (MBP)-tagged wild-type XErp1 but not MBP itself *in vitro* (Fig. 1a). Together with our previous results that XErp1 degradation depends on its prior phosphorylation by Plx1 (ref. 9), this result raised the question of how CaMKII and Plx1 cooperate in response to a calcium signal to target XErp1 for destruction. One attractive model was suggested by recent findings implicating the noncatalytic carboxy terminus of Plks, the so-called polo-box domain (PBD), in both Plk activation and the targeting of these kinases to serine- or threonine-phosphorylated substrates¹². Based on this concept it was tempting to speculate that CaMKII sensitizes XErp1 for degradation by creating a binding site for Plx1. To explore this hypothesis, we examined the XErp1 sequence for sites that could potentially serve as CaMKII-regulated PBD binding sites (RXST/S; ref. 13). Of the ten putative CaMKII phosphorylation sites (RXXS/T) (Fig. 1b) only two sites (RXST¹⁹⁵ and RXST³³⁶) match this optimal consensus. To identify the CaMKII site(s) relevant for calcium-regulated degradation of XErp1 we examined the stability of *in vitro* translated (IVT), ³⁵S-labelled XErp1 mutated at Thr 195 (XErp1^{T195A}) or Thr 336 (XErp1^{T336A}) in calcium-supplemented CSF extract. As shown previously⁹ IVT wild-type XErp1^{WT} was rapidly degraded on anaphase onset whereas XErp1 mutated at its DSGX₃S motif (XErp1^{S33N,S38N}) was not targeted for calcium-induced degradation (Fig. 1c). Notably, IVT XErp1^{T195A}, but not XErp1^{T336A}, remained stable on calcium addition (Fig. 1c), indicating that Thr 195, but not Thr 336, is essential for the cell-cycle-regulated degradation of XErp1. As shown by microscopic analyses and by immunoblotting for cyclin B, all extracts had entered interphase on calcium addition (Fig. 1d, e), confirming that the trace amounts of IVT products added did not block calcium-induced CSF release. Consistent with our previous observations⁹ higher concentrations of XErp1^{WT} and XErp1^{T195A} were capable of blocking calcium-induced CSF release (see Supplementary Fig. S1). Taken together, these data are consistent with the model that CaMKII and Plx1 cooperate to target XErp1 for degradation on a calcium stimulus and that Thr 195 and the DSGX₃S motif are both required for the cell-cycle-regulated destruction of XErp1.

Previous studies have shown that phosphopeptide binding to the Plk-PBD is favoured by a serine residue in position -1 (SpT/pS; ref. 13). The above model thus predicted that Ser 194 might contribute to sensitizing XErp1 for Plx1-dependent degradation. Indeed, ³⁵S-labelled IVT XErp1^{S194A} remained stable in calcium-supplemented CSF extract (Fig. 2a, b), similar to XErp1^{T195A} and XErp1^{S33N,S38N}. To prove that Plx1-binding to XErp1 depends on both Ser 194 and Thr 195, MBP-tagged wild-type protein and appropriate mutants were subjected to far-western experiments. Indeed, phosphorylation of XErp1^{WT} by CaMKII strongly enhanced the interaction between XErp1 and purified full-length Plx1. As expected, wild-type PBD (PBD^{WT}) but not a mutant form of PBD (PBD^{mut}) was capable of binding to phosphorylated XErp1^{WT} (Fig. 2c). In contrast, both the S194A and T195A mutants of XErp1 treated with CaMKII failed to interact significantly with full-length Plx1 or PBD^{WT} (Fig. 2c), indicating that RXST¹⁹⁵ constitutes a CaMKII-regulated Plx1 binding site. Mutation of the DSGX₃S motif did not interfere with the ability of XErp1 to bind to Plx1 or the PBD^{WT} on phosphorylation by CaMKII (Fig. 2c), in line with the expected order of events. Taken together, these data strongly suggest that CaMKII phosphorylation of XErp1 on Thr 195 creates a binding site for Plx1, consistent with the idea that CaMKII sensitizes XErp1 for Plx1-dependent degradation on anaphase onset.

An additional corollary of the proposed model is that PBD docking to XErp1 stimulates the kinase activity of Plx1 towards its substrate XErp1. To test this prediction, we used MBP-XErp1 with or without prior phosphorylation (with unlabelled ATP) by CaMKII to carry out Plx1 kinase assays in the presence of ³²P-labelled ATP. Compared to untreated XErp1^{WT}, XErp1^{WT} pre-phosphorylated by

CaMKII represented a much better *in vitro* substrate for Plx1 (Fig. 2d) consistent with enhanced Plx1 recruitment after CaMKII phosphorylation. To rule out a contribution of CaMKII to the incorporation of ³²P into XErp1, parallel phosphorylation experiments were performed in the absence of Plx1. As shown in Fig. 2d, no labelling of XErp1 with ³²P was observed under these conditions. Notably, pre-treatment of MBP-XErp1^{S194A} or -XErp1^{T195A} did not result in increased phosphorylation by Plx1 (Fig. 2d), confirming the inability of these mutants to provide a docking site for Plx1. Taken together, these data demonstrate that calcium-activated CaMKII converts XErp1 into an efficient Plx1 substrate.

We next wanted to confirm that the activity of CaMKII is essential for the degradation of XErp1 on anaphase onset. However, as the inhibition of CaMKII prevents CSF release, and thus XErp1 degradation, we analysed the stability of IVT XErp1 in extract that was arrested at anaphase by the presence of non-degradable cyclin B (refs 14, 15). These extracts, called 'Δ90 extracts' (refs 14, 15), retain

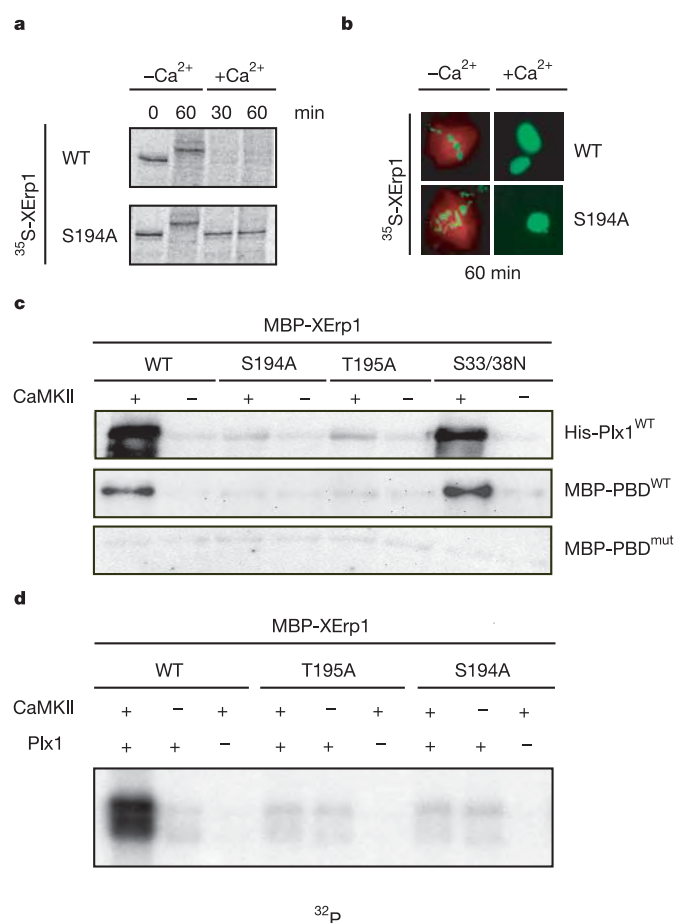


Figure 2 | CaMKII converts XErp1 into an efficient Plx1 substrate.

a, b, Ser 194 is critical for the calcium-induced degradation of XErp1. IVT ³⁵S-XErp1^{WT} or -XErp1^{S194A} was incubated in CSF extract in the presence or absence of calcium and samples were analysed by autoradiography (**a**) and microscopic examination of spindle and DNA morphology (**b**). **c**, Phosphorylation of XErp1 by CaMKII strongly enhances binding of Plx1 to XErp1. The binding of His-Plx1, MBP-PBD^{WT} and MBP-PBD^{mut} to MBP-XErp1 was analysed by far-western ligand blots. MBP-XErp1 proteins were subjected to CaMKII phosphorylation reactions or mock treatments and resolved by SDS-PAGE. Bound Plx1 and PBD were detected using purified anti-Plx1 antibodies. **d**, MBP-XErp1 was subjected to *in vitro* phosphorylation reactions using recombinant Plx1. The incorporation of ³²P was analysed by autoradiography. As indicated, MBP-XErp1 was pre-incubated with activated CaMKII (or subjected to a mock kinase reaction) and unlabelled ATP. To control for incorporation of ³²P by CaMKII, parallel reactions were performed in the absence of Plx1.

CaMKII activity (see Supplementary Fig. S2) which can be modulated without affecting cell cycle progression. As expected, IVT XErp1^{WT} was rapidly degraded in $\Delta 90$ extracts (see Supplementary Fig. S2). As most available specific CaMKII inhibitors are not able to inhibit activated CaMKII but only to prevent CaMKII activation we first suppressed CaMKII activity in $\Delta 90$ extracts by the addition of EGTA (see Supplementary Fig. S2), allowing us to study the effect of a specific CaMKII inhibitor on the stability of XErp1. Consistently, in the absence of an additional calcium stimulus IVT XErp1^{WT} remained stable in EGTA-treated $\Delta 90$ extracts, whereas it was rapidly degraded on calcium addition in a Thr 195-dependent manner (Fig. 3a). Notably, the calcium-induced degradation of XErp1^{WT} could be significantly suppressed by the addition of 300 μ M CaMKII^{281–309} (Fig. 3a), a peptide known to specifically inhibit the calcium-induced activation of CaMKII (ref. 16). CaMKII activity assays demonstrated that CaMKII^{281–309} prevented the calcium-induced reactivation of CaMKII (Fig. 3c). The rapid degradation of IVT securin (Fig. 3a) under all conditions confirmed that the different treatments did not interfere with anaphase arrest. Furthermore, addition of a constitutively active form of CaMKII (CaMKII^{1–290}) induced the rapid degradation of IVT XErp1^{WT} but not of XErp1^{T195A} (Fig. 3b) confirming that CaMKII is the kinase triggering XErp1 degradation on calcium stimulus. In view of an

ongoing debate about a possible contribution of Emi1 to CSF activity¹⁷, we also used the above system to examine the fate of this XErp1-related protein. We found that CaMKII activity is not essential for the degradation of Emi1, as indicated by the fact that IVT Emi1 was efficiently degraded in EGTA-treated $\Delta 90$ extracts even without calcium addition (Fig. 3a), or when the calcium-induced activation of CaMKII was suppressed by the addition of CaMKII^{281–309} (Fig. 3a). Taken together, these data suggest that XErp1, but not Emi1, is the critical target of calcium-activated CaMKII. This conclusion is in line with results showing that Emi1 is highly unstable in CSF-arrested extract^{18,19}, from which CaMKII activity is naturally absent. The observation that Emi1 stability does not seem to be regulated by CaMKII, together with results showing that Emi1 protein is unstable in CSF-arrested egg extract^{18,19}, argues against a critical contribution of Emi1 to CSF activity in *Xenopus* eggs.

The above data suggested that the lack of CaMKII activity accounts for the stability of XErp1 in CSF-arrested *Xenopus* egg extract. Consequently, we reasoned that a mutant XErp1 capable of serving as an efficient Plx1 substrate independently of CaMKII should be targeted for degradation in CSF extract even in the absence of a calcium signal. To test this idea, we converted the CaMKII site (RXST¹⁹⁵LXD) into a consensus site for cyclin dependent kinase 1 (Cdk1; RXST¹⁹⁵PXX) and examined the stability of the ³⁵S-labelled

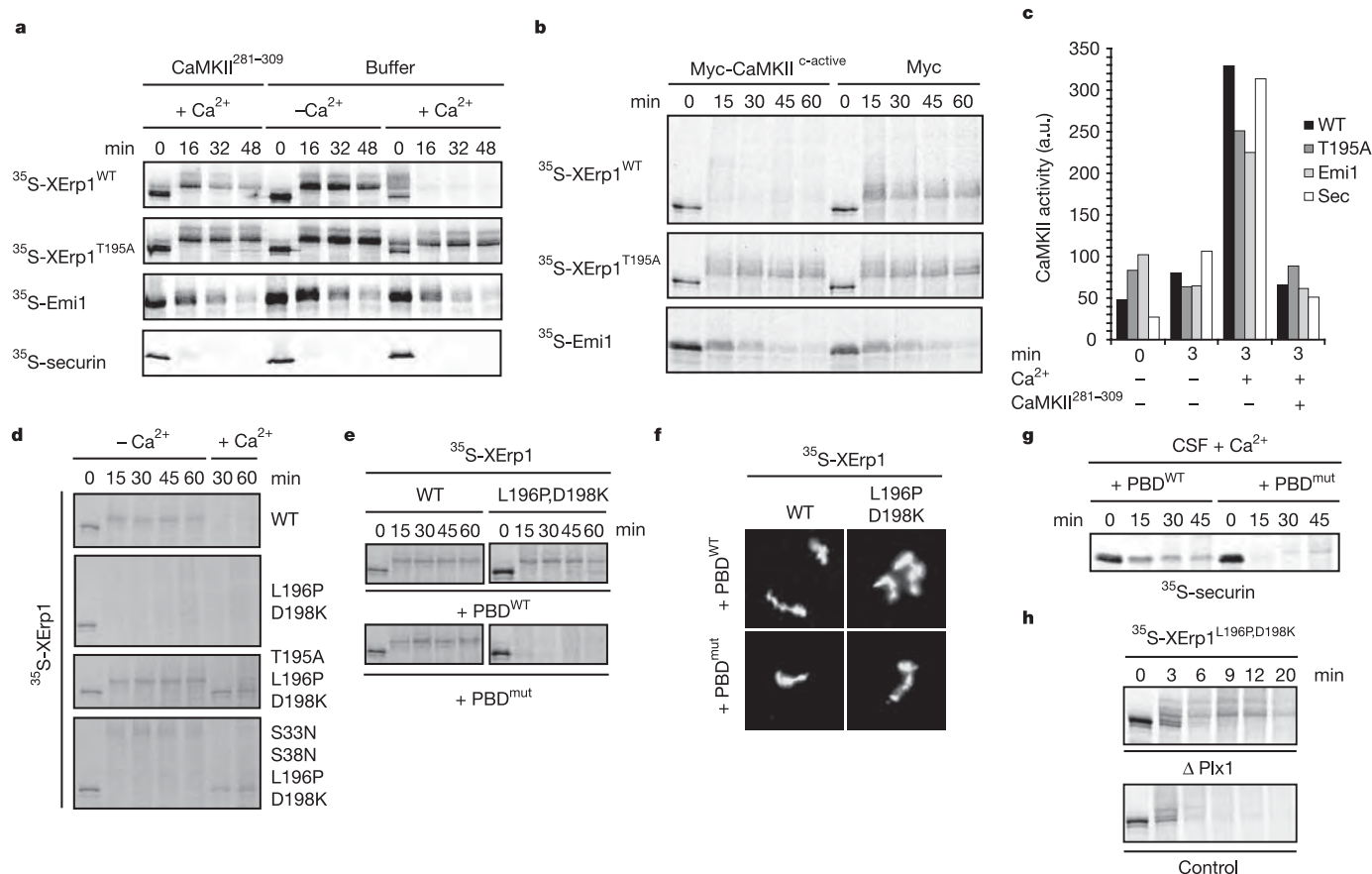


Figure 3 | CaMKII activity is essential for the degradation of XErp1 in *Xenopus* egg extract. **a**, IVT ³⁵S-XErp1^{WT}, -XErp1^{T195A}, -Emi1 or -securin was incubated in $\Delta 90$ extract supplemented with buffer, or calcium in the presence or absence of 300 μ M of the CaMKII inhibitory peptide CaMKII^{281–309}. Samples were analysed by autoradiography. **b**, IVT ³⁵S-XErp1^{WT}, -XErp1^{T195A} or -Emi1 was incubated in $\Delta 90$ extract supplemented with immunopurified IVT constitutively active CaMKII^{c-active} or immunopurified mock IVT and analysed as in **a**. **c**, Samples from extracts shown in **a** were analysed for CaMKII activity. **d**, Cdk1 can take on CaMKII's function in sensitizing XErp1 for Plx1-dependent degradation. IVT

³⁵S-XErp1^{WT}, -XErp1^{L196P,D198K}, -XErp1^{T195A,L196P,D198K} or -XErp1^{S33N,S38N,L196P,D198K} was incubated in CSF extract and samples were analysed by autoradiography. **e–h** Plx1 activity is required for the degradation of XErp1^{L196P,D198K} in CSF extract. **e**, **f**, Samples of CSF extracts supplemented with MBP-PBD^{WT} or MBP-PBD^{mut} were analysed (**e**) as in **d**, and assayed for DNA morphology (**f**). CSF extracts treated as in **e** were supplemented with calcium and assayed for the stability of IVT securin (**g**). Plx1-depleted (Δ Plx1) or mock depleted (control) CSF extract were treated as in **d** (**h**).

IVT product in CSF extracts. As Cdk1 is active in CSF extract and able to create docking sites for Plx1-PBDs (ref. 13), the prediction was that Cdk1-mediated phosphorylation of this mutant XErp1 would result in the docking of Plx1, the phosphorylation of the amino-terminal phospho-degron and the destruction of the protein. Indeed, whereas IVT XErp1^{WT} remained stable in CSF extract, XErp1^{L196P,D198K} was rapidly degraded even in the absence of calcium (Fig. 3d), indicating that Cdk1 or any other proline-directed kinase active in CSF extract could substitute for CaMKII and sensitize this XErp1 mutant for Plx1-dependent degradation. As predicted by our model, the degradation of these XErp1 forms in CSF extract could be prevented by mutating either critical residues within the CDK1 consensus site (XErp1^{T195A,L196P,D198K}) or the DSGX₃S³⁸ motif (XErp1^{S33N,S38N,L196P,D198K}) (Fig. 3d). To demonstrate that the activity of Plx1 was essential for the degradation of XErp1^{L196P,D198K}, we examined the stability of IVT XErp1^{L196P,D198K} in CSF extract supplemented with an excess of MBP-PBD^{WT}, which we have shown to exert a dominant-negative effect on Plx1 function⁹. As expected, the addition of MBP-PBD^{WT} resulted in a significant stabilization of IVT XErp1^{L196P,D198K} in CSF extract, as compared to a control-treated extract (Fig. 3e) but had no effect on the stability of IVT XErp1^{WT} (Fig. 3e). Analyses of the chromatin structures revealed that the extracts remained CSF-arrested under all conditions (Fig. 3f). The dominant-negative effect of MBP-PBD on Plx1 function was confirmed in parallel experiments in which we examined the stability of IVT securin on calcium addition. PBD^{WT}-supplemented CSF extract, but not PBD^{mut}-treated extract, failed to exit meiosis on calcium treatment, as indicated by stable IVT securin (Fig. 3g). Consistently, MBP-PBD^{WT} but not PBD^{mut} prevented the calcium-induced degradation of endogenous XErp1 (see Supplementary Fig. S3). Finally, a similar stabilization of IVT XErp1^{L196P,D198K} was observed in CSF extract immunodepleted of Plx1 but not in mock-depleted extract (Fig. 3h). Taken together, these data demonstrate that the creation of a Plx1-PBD docking site on XErp1, an event normally brought about by calcium-activated CaMKII, defines the timing of Plx1-dependent degradation of XErp1.

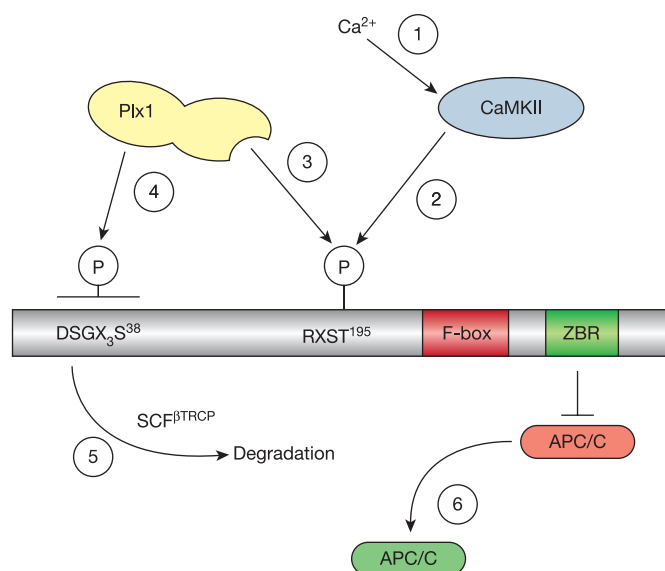


Figure 4 | Model of how calcium triggers release from CSF arrest. (1) A transient rise in free calcium activates CaMKII. (2) Activated CaMKII phosphorylates XErp1 at Thr 195, thereby creating a docking site for Plx1. (3) Plx1 binds to XErp1 via its PBD. (4) On binding to its substrate, Plx1 becomes activated and phosphorylates XErp1 at Ser 33/38 of the DSGX₃S³⁸ motif. (5) Phosphorylated DSGX₃S³⁸ is recognized by SCF^{βTRCP} leading to the destruction of XErp1. (6) XErp1 degradation leads to APC/C activation.

In conclusion, our study identifies the critical substrate of the calcium-activated CaMKII that triggers exit from meiosis II in response to fertilization. Specifically, our data demonstrate that CaMKII phosphorylates the APC/C inhibitor XErp1, thereby acting as a novel priming kinase for the recruitment of Plx1 (Fig. 4). The bound Plx1, previously shown to be essential for causing release from CSF arrest¹¹, then triggers the degradation of XErp1 by phosphorylation of a DSGX₃S³⁸ degron. Our findings thus explain how a calcium signal prompts the exit from meiosis through the spatiotemporal integration of the action of two key kinases, CaMKII and Plx1, both converging onto XErp1, a critical inhibitor of APC/C.

METHODS

Plasmids, proteins and antibodies. XErp1, Emi1 and securin constructs were as described⁹. Site-directed mutagenesis was performed using the QuikChange kit (Stratagene). MBP-tagged proteins were purified following published protocols⁹. Purification of His-tagged Plx1 from SF9 cells was performed as described¹¹. *In-vitro* translation experiments using ³⁵S-labelled methionine were performed according to manufacturer's protocol (Promega). Full-length Plx1 purified from SF-9 cells was used to generate rabbit antibodies to Plx1. MBP-PBD^{WT} and MBP-PBD^{mut} (W408F, H532A and K534A) were used as described previously⁹.

In vitro kinase assays. CaMKII was purchased from New England Biolabs. *In vitro* CaMKII assays were performed at 30 °C in kinase reaction buffer containing 45 U of activated CaMKII, 0.1–1 mM ATP, 4 μCi [γ-³²P]ATP and 200 ng of purified MBP-tagged XErp1 protein. Samples were taken at the indicated time points. CaMKII activity assays were performed based on published protocols⁸ using Autocamtide-2 (New England Biolabs) as substrate. Plx1 kinase assays were performed for 5 min at 30 °C using kinase reaction buffer supplemented with 4 μCi [γ-³²P]ATP and MBP-tagged XErp1 proteins which had previously been incubated with CaMKII or buffer (control) in the presence of 1 mM ATP.

Far-western ligand blots. Far-western ligand blots were performed in TBS supplemented with 0.1% Tween-20 and 5% wt/vol skim milk powder. For each assay 500 ng MBP-XErp1, treated with CaMKII or buffer (control), was subjected to SDS polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to membranes. These were then incubated overnight at 4 °C with 2 μg ml⁻¹ of MBP-tagged PBD⁹ or His-Plx1. Bound protein was detected using affinity-purified rabbit antibodies to Plx1.

Xenopus extracts. *Xenopus* CSF egg extracts were prepared as described previously²⁰. CSF release was induced by adding 1 mM CaCl₂ to the extract. DNA and spindle morphology were examined as described previously²¹. CSF-released extract was arrested at anaphase by the addition of non-degradable cyclin B (Δ90 extract) as described^{14,15}. The Δ90 extract was treated with 300 μM EGTA and CaMKII was re-activated by the addition of 600 μM calcium. Where indicated, MBP-PBD^{WT} or MBP-PBD^{mut} was added to the extract to a final concentration of 400 μg ml⁻¹. For some experiments extract was treated with cycloheximide. The immunodepletion of Plx1 was performed as described²².

Received 19 May; accepted 5 August 2005.

Published online 28 August 2005.

1. Tunquist, B. J. & Maller, J. L. Under arrest: cytotostatic factor (CSF)-mediated metaphase arrest in vertebrate eggs. *Genes Dev.* 17, 683–710 (2003).
2. Masui, Y. & Markert, C. L. Cytoplasmic control of nuclear behaviour during meiotic maturation of frog oocytes. *J. Exp. Zool.* 177, 129–145 (1971).
3. King, R. W. *et al.* A 20S complex containing CDC27 and CDC15 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell* 81, 279–288 (1995).
4. Sudakin, V. *et al.* The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol. Biol. Cell* 6, 185–197 (1995).
5. Salah, S. M. & Nasmyth, K. Destruction of the securin Pds1p occurs at the onset of anaphase during both meiotic divisions in yeast. *Chromosoma* 109, 27–34 (2000).
6. Busa, W. B. & Nuccitelli, R. An elevated free cytosolic Ca²⁺ wave follows fertilization in eggs of the frog, *Xenopus laevis*. *J. Cell Biol.* 100, 1325–1329 (1985).
7. Lorca, T. *et al.* Degradation of the proto-oncogene product p39mos is not necessary for cyclin proteolysis and exit from meiotic metaphase: requirement for a Ca(2+)-calmodulin dependent event. *EMBO J.* 10, 2087–2093 (1991).
8. Lorca, T. *et al.* Calmodulin-dependent protein kinase II mediates inactivation of MPF and CSF upon fertilization of *Xenopus* eggs. *Nature* 366, 270–273 (1993).
9. Schmidt, A. *et al.* *Xenopus* polo-like kinase Plx1 regulates XErp1, a novel inhibitor of APC/C activity. *Genes Dev.* 19, 502–513 (2005).
10. Fuchs, S. Y., Spiegelman, V. S. & Kumar, K. G. The many faces of beta-TrCP E3 ubiquitin ligases: reflections in the magic mirror of cancer. *Oncogene* 23, 2028–2036 (2004).

11. Descombes, P. & Nigg, E. A. The polo-like kinase Plx1 is required for M phase exit and destruction of mitotic regulators in *Xenopus* egg extracts. *EMBO J.* **17**, 1328–1335 (1998).
12. Elia, A. E., Cantley, L. C. & Yaffe, M. B. Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates. *Science* **299**, 1228–1231 (2003).
13. Elia, A. E. *et al.* The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell* **115**, 83–95 (2003).
14. Holloway, S. L., Glotzer, M., King, R. W. & Murray, A. W. Anaphase is initiated by proteolysis rather than by the inactivation of maturation-promoting factor. *Cell* **73**, 13393–13402 (1993).
15. Stemmann, O., Zou, H., Gerber, S. A., Gygi, S. P. & Kirschner, M. W. Dual inhibition of sister chromatid separation at metaphase. *Cell* **107**, 715–726 (2001).
16. Colbran, R. J., Fong, Y. L., Schworer, C. M. & Soderling, T. R. Regulatory interactions of the calmodulin-binding, inhibitory, and autophosphorylation domains of Ca^{2+} /calmodulin-dependent protein kinase II. *J. Biol. Chem.* **263**, 18145–18151 (1988).
17. Zachariae, W. Emi1 and Erp1: who can stop these eggs? *Dev. Cell* **8**, 301–303 (2005).
18. Ohsumi, K., Koyanagi, A., Yamamoto, T. M., Gotoh, T. & Kishimoto, T. Emi1-mediated M-phase arrest in *Xenopus* eggs is distinct from cytostatic factor arrest. *Proc. Natl Acad. Sci. USA* **101**, 12531–12536 (2004).
19. Tung, J. J. *et al.* A role for the anaphase-promoting complex inhibitor Emi2/XErp1, a homolog of early mitotic inhibitor 1, in cytostatic factor arrest of *Xenopus* eggs. *Proc. Natl Acad. Sci. USA* **102**, 4318–4323 (2005).
20. Murray, A. W. Cell cycle extracts. *Methods Cell Biol.* **36**, 581–605 (1991).
21. Desai, A., Murray, A., Mitchison, T. J. & Walczak, C. E. The use of *Xenopus* egg extracts to study mitotic spindle assembly and function *in vitro*. *Methods Cell Biol.* **61**, 385–412 (1999).
22. Funabiki, H. & Murray, A. W. The *Xenopus* chromokinesin Xkid is essential for metaphase chromosome alignment and must be degraded to allow anaphase chromosome movement. *Cell* **102**, 411–424 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank O. Stemmann and F. Barr for reagents and discussions during this work. We are grateful to I. Gorr and D. Boos for advice on $\Delta 90$ experiments, S. Hümmer for continuous support and R. Neef for assistance with the far-western analyses.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.U.M. (mayer@biochem.mpg.de).

Solution structure of a protein denatured state and folding intermediate

T. L. Religa¹, J. S. Markson¹, U. Mayor^{1†}, S. M. V. Freund¹ & A. R. Fersht¹

The most controversial area in protein folding concerns its earliest stages. Questions such as whether there are genuine folding intermediates, and whether the events at the earliest stages are just rearrangements of the denatured state¹ or progress from populated transition states², remain unresolved. The problem is that there is a lack of experimental high-resolution structural information about early folding intermediates and denatured states under conditions that favour folding because competent states spontaneously fold rapidly. Here we have solved directly the solution structure of a true denatured state by nuclear magnetic resonance under conditions that would normally favour folding, and directly studied its equilibrium and kinetic behaviour. We engineered a mutant of *Drosophila melanogaster* Engrailed homeodomain that folds and unfolds reversibly just by changing ionic strength. At high ionic strength, the mutant L16A is an ultra-fast folding native protein, just like the wild-type protein; however, at physiological ionic strength it is denatured. The denatured state is a well-ordered folding intermediate, poised to fold by docking helices and breaking some non-native interactions. It unfolds relatively progressively with increasingly denaturing conditions, and so superficially resembles a denatured state with properties that vary with conditions. Such ill-defined unfolding is a common feature of early folding intermediate states and accounts for why there are so many controversies about intermediates versus compact denatured states in protein folding.

Drosophila melanogaster Engrailed homeodomain (En-HD) is a 61-residue, three-helix bundle protein with helices spanning residues 10–22 (H1), 28–37 (H2) and 42–56 (H3)³. It folds on the micro-second timescale via a folding intermediate, and its pathway is well studied by experiment and simulation⁴. We have engineered En-HD to be denatured under conditions that favour folding by introducing the mutation L16A (ref. 5). L16 is a highly conserved, fully buried residue located in the middle of H1. The L16A mutation removes a number of local interactions and numerous long-range interactions with residues in H2, the turn between H2 and H3, but mainly with H3. Φ -analysis and simulation show that the core becomes disrupted in this region in the transition state for unfolding, but the helical secondary structure of H1 remains intact, as does that of H2 and H3 and the turn connecting them⁴. The L16A mutation is likely to destabilize the native structure substantially but not folding intermediates. All evidence so far shows that L16A is not only a genuine model for the denatured state of En-HD under folding conditions but is also a denatured state in its own right. The mutation acts to tip the energetic balance between native and denatured states so that L16A is denatured at physiological ionic strength⁵. Addition of salt stabilizes the protein and screens charge–charge repulsions so that it refolds⁵.

It is important first to define the term ‘denatured state’ explicitly. We adopt the convention that the maximally unfolded state of a protein, in which the backbone NH groups have little protection

against ¹H/²H-exchange, is U. The denatured state, D, is the lowest energy non-native state under a defined set of conditions⁶. Under physiological conditions that favour folding, the lowest energy denatured state, D_{phys}, can be a folding intermediate⁷. Native state ¹H/²H-exchange kinetics has shown, for example, that the denatured state of En-HD has protected helices H1, H2 and H3, and the U state is at 1 kcal mol^{−1} higher energy⁴.

Chemical shifts⁸, nuclear Overhauser effects (NOEs)^{9,10}, paramagnetic relaxation^{11,12} and residual dipolar couplings (RDCs)¹³ are used to infer the structural properties of the denatured state ensemble and calculate its three-dimensional structure¹⁴. We solved the structure of the denatured En-HD mutant L16A at physiological ionic strength (~150 mM) using NOEs and classical nuclear magnetic resonance (NMR) procedures¹⁵ (Protein Data Bank code 1ZTR; Fig. 1a). Over 800 intra-residue NOEs were identified. Over 90 of them were long-range NOEs, with most of them occurring between H2 and H3 (Supplementary Table 1). Experiments on deuterated samples identified only one significant long-range contact—between I45 HN and S35 HN. We found no NOEs between H1 and H2 or H3. The detected sequential NOEs together with H α , C α and C β chemical shifts (see Fig. 1b, c) implied that the protein displays helical properties at least between residues 10–25, 28–38 and 42–53. These regions differ from the wild type, which is not helical around residues 22–28 and has a longer H3. The structure of the region spanning H2 and H3 and the connecting loop is as well defined as many NMR structures of native proteins (backbone root mean square deviation (r.m.s.d.) of 0.87 Å) and similar to that of the wild-type protein. But, although H1 was found to be helical, the angle between it and the rest of the protein is only approximate because of the lack of long-range NOEs between them. It may well be that the angle between H1 and the rest of the protein does indeed fluctuate.

H1 is longer in the L16A mutant than in the wild-type protein by at least one turn. Numerous hydrophobic residues become solvent exposed and form contacts not observed in the wild type. Residues F8 and L13 of the L16A mutant, which interact with H3, turn residues and L16 in the wild-type protein, have NOEs between each other which are unambiguously different from those in the wild-type protein. F20, which interacts mainly with the carboxy terminus of H3 in the wild type, bends back towards H1 in the L16A mutant, showing numerous non-wild-type NOEs with residues close in sequence. Y25, which is solvent-exposed in the wild type, where it forms long-range interactions with R53, has numerous non-wild-type NOEs to residues in the wild-type loop region. T27 H γ_2 , fully exposed in the wild type, has an unambiguous NOE with the F49 aromatic ring in the L16A mutant. Significant portions of W48 and F49 are buried by interactions with H1 in the wild type. In the L16A mutant, both residues bend back to the face of H3 and become partially buried in the H2-loop-H3 hydrophobic core.

¹MRC Centre for Protein Engineering and Cambridge University Chemical Laboratories, MRC Centre, Hills Road, Cambridge CB2 2QH, UK. [†]Present address: Wellcome Trust Cancer Research UK Gurdon Institute, Tennis Court Road, Cambridge CB2 1QN, UK.

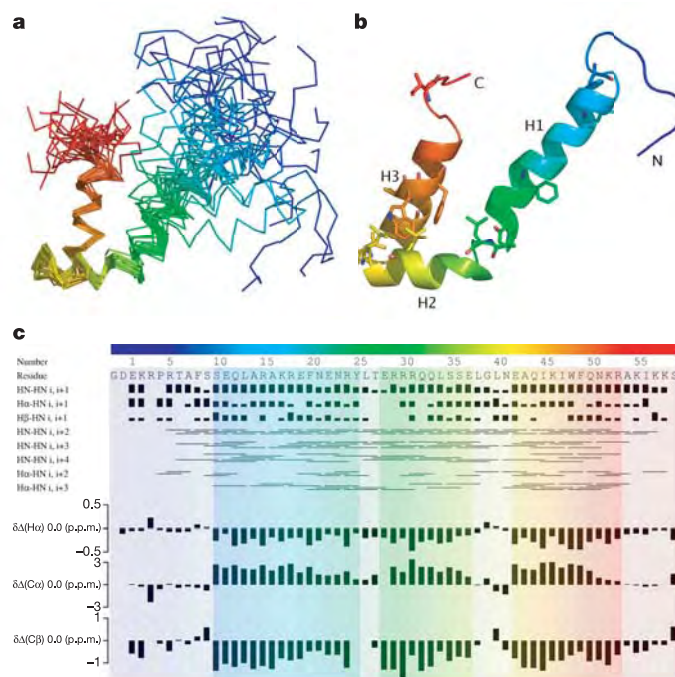


Figure 1 | Structure of En-HD L16A at low ionic strength and summary of sequential NOEs and secondary chemical shift values (Protein Data Bank code 1ZTR). **a**, Bundle of the 25 lowest-energy conformers with backbone atoms N, C α and C' of residues 28–53 superimposed for a minimal r.m.s.d. The different conformations of residues 1–28 are a consequence of the lack

of a significant number of long-range NOEs between them and residues 28–53. **b**, Lowest energy structure with location of hydrophobic side chains. **c**, Sequential NOEs and H α , C α and C β chemical shift deviations. Negative, positive and negative deviations from random-coil values for H α , C α and C β chemical shifts, respectively, indicate helical structure.

We examined the structure of the backbone by HN-N RDC measurements using acrylamide and Pf1 phage as alignment media (Supplementary Information). The RDCs were fully consistent with the NOEs. The structures encompassing helices 2 and 3 superposed well on those derived from NOEs, and the orientation of H1 in the RDC structures was well within the envelope of the NOE structures. We analysed ^{15}N relaxation data for the L16A mutant with the reduced spectral density approach at 25 °C (refs 16–18) to measure the degree of mobility along its chain. The results show that it is a well-ordered protein (Fig. 2; see also Supplementary Information).

The structure of L16A is fully consistent with previous experimental data, including its circular dichroism, fluorescence spectra and dimensions and shape from small angle X-ray scattering, and also with the native-state hydrogen/deuterium (H/D)-exchange pattern^{4,5} and multiple molecular dynamics simulations of the unfolding of wild-type En-HD⁴. The structure of L16A is also perfectly consistent with the Φ -analysis of the structure of the transition state for unfolding of En-HD⁴, with the helix-turn-helix motif around H2 and H3 being fully formed in the transition state ($\Phi \approx 1$), as is the secondary structure in H1, with the core between H1 and the rest of the protein coming apart.

Having found that the L16A mutant is a good model for the denatured state of En-HD and is a well-structured intermediate that is poised for folding, we investigated the L16A mutant as a genuine denatured state and folding intermediate in its own right in equilibrium with its native state. L16A went from having little detectable native tertiary structure at low ionic strength (145 mM) (ref. 5) to being $\sim 98\%$ folded at 3.5 M NaCl, with a free energy of denaturation of 2.3 kcal mol $^{-1}$ (Fig. 3a, c). Denaturation of L16A in 0.1 M NaCl had an apparently non-cooperative transition for melting of the helical structure⁵, monitored by circular dichroism. However, in 4 M NaCl, L16A melted with increasing temperature according to a classical cooperative two-state transition that paralleled the thermal denaturation of the wild-type protein (Fig. 3c). The heteronuclear single quantum correlation (HSQC) NMR spectrum of the L16A

mutant in 4 M NaCl had a chemical shift dispersion almost as high as that of wild type (data not shown). A plot of the circular dichroism signal at 222 nm or of tryptophan fluorescence against the logarithm of the ionic strength showed a sharp cooperative transition for formation of native structure, with a mid-point of about 1 M (Fig. 3d). Comparison of the stability of L16A and wild type at high ionic strength implies that the denatured state of L16A is favoured by about 2 kcal mol $^{-1}$ at the ionic strength of 145 mM used in NMR studies. Wild-type En-HD folds according to apparent two-step kinetics, with an unassigned ultra-fast phase at the 1- μs region and a very fast phase at the 10- μs region that has been assigned by NMR line broadening to correspond to the formation of native structure⁴. At low ionic strength, L16A exhibits the 1- μs phase only⁴. We saw (Fig. 4) both slow and fast phases for L16A on addition of 500 mM NaCl, and mainly the slower phase at 2 M NaCl. The data agree well with those of wild-type En-HD at low salt, with the unfolding limb of the plot being observable at a lower temperature.

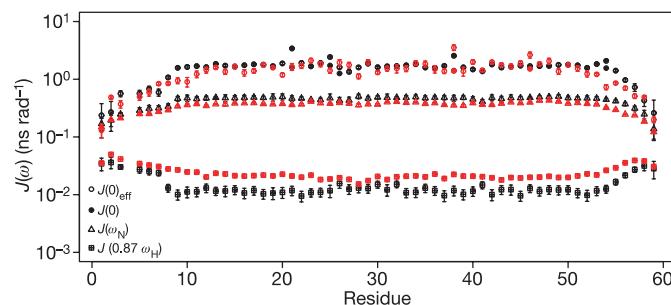


Figure 2 | Backbone dynamics of wild-type and En-HD L16A protein. Reduced spectral densities for En-HD L16A (red) and wild type (black) at 25 °C. $J(0)$ relaxation rates compensated for chemical exchange using η/R_2 (ref. 23) are shown with filled circles.

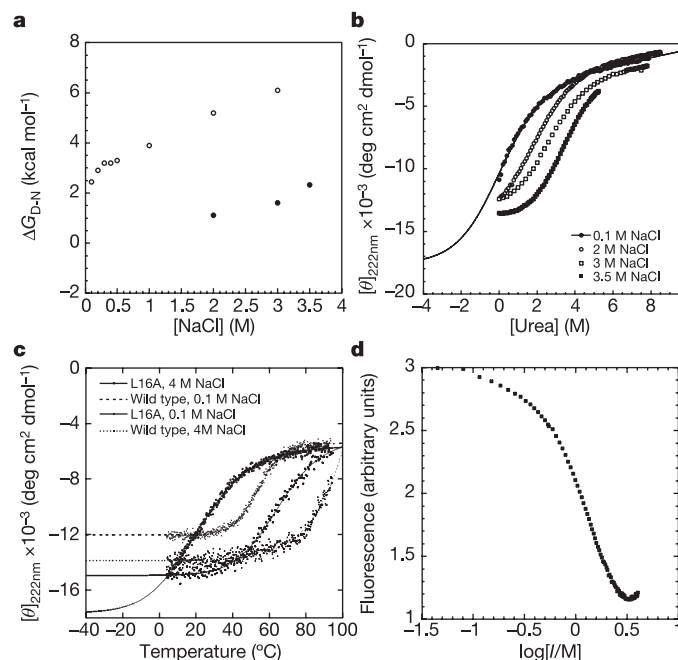


Figure 3 | Effects of salt on the properties of wild-type En-HD and L16A, and the difficulties of distinguishing between cooperative and non-cooperative transitions. **a**, Free energies of denaturation, ΔG_{D-N} , for wild-type and L16A at 25 °C, as a function of NaCl concentration. Open circles, wild-type ΔG_{D-N} calculated from urea denaturation experiments; filled circles, L16A ΔG_{D-N} calculated from urea denaturation. **b**, Urea denaturation of L16A monitored by circular dichroism in the presence of 0.1, 2, 3, or 3.5 M NaCl, at 25 °C. Data were fitted to the standard two-state denaturation equation. At 0.1 M NaCl, we would need experimental data at negative concentrations of urea to fit the curve with any confidence—we used the limiting value of the circular dichroism signal from **c** and assumed a zero baseline slope in the absence of the necessary data to give a value of m of 0.5 with $[\text{urea}]_{50\%}$ of 0 M (compare with $m = 0.71$ for wild type or the L16A mutant at high salt). **c**, Thermal denaturation of wild-type En-HD and L16A, monitored by circular dichroism at various salt concentrations. We can fit the data at 0.1 M to a cooperative transition by assuming that the baseline at low temperature has zero slope, to give a T_m of 21 °C and ΔH_{D-N} of 12 kcal mol⁻¹. We would need to measure down to -20 to -40 °C for a confident fit. **d**, Titration of En-HD L16A with NaCl, at 25 °C, with fluorescence amplitude plotted against $\log[I]$ (ref. 24), where I is the ionic strength.

At 25 °C and 500 mM salt, where it is ~10% folded, the L16A mutant unfolded at 60,000 s⁻¹ and folded at 6,000 s⁻¹. In 2 M salt, refolding increased to ~20,000 s⁻¹ and unfolding decreased to ~3,000 s⁻¹. L16A is thus a fast folding protein with rate constants comparable to those of wild type. The L16A mutant at low ionic strength is not just a model for the denatured state of En-HD but is a true denatured state in equilibrium with its native state in a conventional folding reaction for which the equilibrium position changes with ionic strength.

We can now use the properties of L16A to shed light on controversies about intermediates and denatured states in general because the denatured state of L16A is a well-defined and chemically reasonable high-resolution structure. One conventional test for a folded structure when it cannot be directly characterized is the observation of a cooperative folding transition. Larger proteins that have been traditionally studied usually display sharp folding transitions. But, with very small proteins the transitions become very broad. For example, a protein that has an enthalpy of denaturation (ΔH_{D-N}) of only 12 kcal mol⁻¹ would require a temperature range of about the melting temperature (T_m) ± 60 –70 °C to go from 98–99% native to 98–99% denatured; such a low and spread-out value of ΔH_{D-N} would be difficult to detect by calorimetry. Similarly, an m (that is, the rate of change of free energy of denaturation with concentration of urea, $\partial \Delta G_{D-N} / \partial [\text{urea}]$) of 0.5 would require a

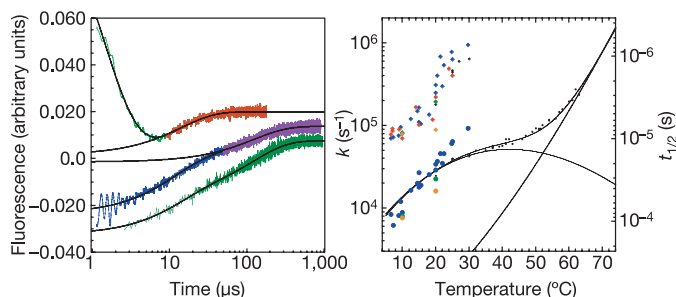


Figure 4 | Relaxation kinetics of wild-type En-HD and the L16A mutant at various concentrations of NaCl as a function of temperature at 50 mM NaAc, pH 5.7. The left panel shows time courses for the relaxation kinetics of L16A. The red curve is for 100 mM NaCl; the green at the beginning is the response time of the T-jump machine, which is long at this ionic strength. The theoretical curves are for a single exponential through the folding relaxation and a double to include the response time. Blue and purple are fast and slow phases at 500 mM NaCl. Dark green is the trace with 1 M NaCl. The fits in the latter two are double exponentials for the actual two phases. The right panel shows relaxation rate constants versus temperature. Diamonds, fast phase; circles, slow phase; red, En-HD L16A in 100 mM NaCl; blue, En-HD L16A in 500 mM NaCl; dark green, En-HD L16A in 1 M NaCl; orange, En-HD L16A in 2 M NaCl; black, En-HD wild type in 50 mM NaAc pH 5.7. The thin black lines are the deconvoluted folding and unfolding rate constants for the slow folding phase of the wild-type protein.

denaturant concentration range of 9–11 M to go from 98–99% native to 98–99% denatured. In practice, a weakly stable small protein can be studied over only such a limited range of temperatures and denaturant such that it is not possible to determine with confidence whether a transition is cooperative and to determine T_m values with precision because the baselines of curves are inaccessible. The folding transition of the L16A mutant under physiological conditions is a typical ill-defined example, and we have equivocated on whether it is a discrete state because it was difficult to argue whether the transition was cooperative or whether we were just fitting the middle regions of a non-cooperative transition with a mixture of states⁵. Now that we have characterized L16A as being folded, we have fitted its denaturation at physiological ionic strength to cooperative denaturation transitions to give a T_m of ~33–36 °C and ΔH_{D-N} of ~13 kcal mol⁻¹ (Fig. 3). In particular, the thermal titration of more than 30 residues monitored by NMR fits the middle of sigmoidal transition curves (data not shown; $T_m = 36.3 \pm 4.3$ °C and ΔH_{D-N} of 13.4 ± 3.4 (±s.d.)), the caveat being that there is a small spread in T_m that may be accounted for by the uncertainties in not knowing the endpoints or slopes of baselines of the transitions. It is very rare to find a folding intermediate as well ordered and stable as that of the L16A mutant. Most other examples of intermediates cannot be isolated for direct study and will have even lower energy transitions that will require impossibly large temperature ranges to distinguish between cooperative and non-cooperative transitions. It is, accordingly, nearly impossible in most cases to distinguish experimentally between discrete intermediates and denatured states of varying compaction.

METHODS

NMR methods. NMR experiments were performed in 50 mM sodium acetate-d₃ buffer, pH 5.7, 100 mM NaCl and at temperatures of 25 and 5 °C on Bruker DRX500 and Avance 800 spectrometers equipped with 5-mm triple-resonance probes and single-axis gradients. Protein concentrations were 600 μM for nuclear Overhauser effect spectroscopy (NOESY) experiments and 100 μM for the RDC measurements using protein prepared as described in the Supplementary Information. There were no changes in chemical shifts and NOESY spectra within this range. Resonance assignments and NOESY spectra were obtained at 5 °C. At this temperature, the amide chemical shift dispersion for En-HD L16A was 1.4 p.p.m. with most of the residues located within a 0.8-p.p.m. window. Backbone assignments were obtained with standard triple-resonance experiments¹⁵, and two-dimensional constant-time-[¹³C,¹H]-HSQC and three-dimensional

[^{13}C , ^{15}N , ^1H]-(H)CC(CO)NH-TOCSY (TOCSY, total correlated spectroscopy) spectra enabled a 95% complete side-chain assignment, including stereo-specific assignments of the methyl groups. Structural NOEs were obtained from two-dimensional [^1H , ^1H]-NOESY as well as three-dimensional [^{15}N , ^1H , ^1H]-NOESY and [^{13}C , ^1H , ^1H]-NOESY spectra. The mixing time in all experiments was 150 ms to minimize spin diffusion. A 100% deuterated ^{15}N -labelled sample was used to record three-dimensional [^{15}N , ^{15}N , ^1H]-HSQC-NOESY-HSQC and two-dimensional [^1H , ^1H]-NOESY experiments with a mixing time of 600 ms^{10,19}. These experiments, optimized for long-range (>5 Å) contacts, confirmed weak NOEs obtained with 150 ms mixing time but detected no additional long-range interactions between H1 and H2/H3. A 600-ms three-dimensional [^{15}N , ^1H , ^1H]-NOESY and two-dimensional [^1H , ^1H]-NOESY were used to confirm amide to aromatic side-chain interactions in a 100% deuterated, Phe, Tyr back-protonated sample. Cross-peaks from the 150-ms NOESY spectra were divided into three bins with cutoffs at 2.7, 3.6 and 5.5 Å based on their volumes. Amide–amide NOEs from the 600-ms three-dimensional [^{15}N , ^{15}N , ^1H]-HSQC-NOESY-HSQC and amide–aromatic NOEs from the NOESY spectra were divided into four bins with cutoffs at 2.7, 3.6, 4.7 and 7.2 Å. On the basis of these restraints, 50 structures were calculated using the CNS²⁰ package and the 25 lowest energy structures are presented here.

We checked whether assuming a single entity for the L16A mutant affected the structure determination. We applied ensemble averaging as implemented in CNS^{21,22} to account for possible ensemble solution to our NOEs. We obtained the same structure but with lower precision, decreasing to 2.1 Å for the H2/H3 fragment. The NMR data are consistent with a single entity, but, as is true for all structures, cannot exclude an ensemble of rapidly interconverting structures.

^{15}N backbone amide relaxation data for En-HD L16A and wild type were acquired and analysed as described previously⁵.

Thermodynamic and kinetic characterization. The standard buffer was 100 mM NaCl, 50 mM NaAc pH 5.7 ($I = 145 \text{ mM}$), with I increased by addition of NaCl. Folding equilibria and kinetics of L16A and wild-type protein were studied as described^{4,5}. Urea denaturations and salt titrations were performed at 25 °C using an Aviv 202-SF instrument, with fluorescence recorded through a 320-nm long-pass filter and a protein concentration of $\sim 8 \mu\text{M}$. Thermal denaturations were performed on a Jasco J-720 spectrometer with a heating rate of 60–80 °C per hour and a protein concentration generally ~ 10 – $20 \mu\text{M}$. Kinetic experiments were performed on a DIA-RT T-jump (DiaLog) or a TJ-64 T-jump (Hi-Tech) electrical-discharge temperature-jump instrument. Protein concentrations were 15–300 μM in acetate or HEPES⁴.

Received 11 March; accepted 19 July 2005.

- Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. & Sosnick, T. R. Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371 (2002).
- Yang, W. Y. & Gruebele, M. Folding at the speed limit. *Nature* **423**, 193–197 (2003).
- Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L. & Pabo, C. O. Structural studies of the engrailed homeodomain. *Protein Sci.* **3**, 1779–1787 (1994).
- Mayor, U. *et al.* The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* **421**, 863–867 (2003).
- Mayor, U., Grossmann, J. G., Foster, N. W., Freund, S. M. & Fersht, A. R. The denatured state of Engrailed Homeodomain under denaturing and native conditions. *J. Mol. Biol.* **333**, 977–991 (2003).
- Dill, K. A. & Shortle, D. Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795–825 (1991).
- Oliveberg, M. & Fersht, A. R. Thermodynamics of transient conformations in the folding pathway of barnase: Reorganization of the folding intermediate at low pH. *Biochemistry* **35**, 2738–2749 (1996).
- Korzhnev, D. M. *et al.* Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590 (2004).
- Neri, D., Billeter, M., Wider, G. & Wuthrich, K. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* **257**, 1559–1563 (1992).
- Mok, Y. K., Kay, C. M., Kay, L. E. & Forman-Kay, J. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**, 619–638 (1999).
- Gillespie, J. R. & Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **268**, 158–159 (1997).
- Gillespie, J. R. & Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **268**, 170–184 (1997).
- Mohana-Borges, R., Goto, N. K., Kroon, G. J., Dyson, H. J. & Wright, P. E. Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.* **340**, 1131–1142 (2004).
- Dyson, H. J. & Wright, P. E. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods Enzymol.* **339**, 258–270 (2001).
- Wuthrich, K. *NMR of Proteins and Nucleic Acids* (John Wiley & Sons, New York, 1986).
- Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A. & Kay, L. E. Spectral density function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR* **6**, 153–162 (1995).
- Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry* **34**, 868–878 (1995).
- Ishima, R. & Nagayama, K. Protein backbone dynamics revealed by quasi spectral density function analysis of amide N-15 nuclei. *Biochemistry* **34**, 3162–3171 (1995).
- Koharudin, L. M. I., Bonvin, A., Kaptein, R. & Boelens, R. Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination. *J. Mag. Res.* **163**, 228–235 (2003).
- Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for molecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Bonvin, A. M. J. J. & Brunger, A. T. Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol.* **250**, 80–93 (1995).
- Bonvin, A. M. J. J. & Brunger, A. T. Do NOE distances contain enough information to assess the relative populations of multi-conformer structures? *J. Biomol. NMR* **7**, 72–76 (1996).
- Fushman, D., Tjandra, N. & Cowburn, D. Direct measurement of N-15 chemical shift anisotropy in solution. *J. Am. Chem. Soc.* **120**, 10947–10952 (1998).
- Rios, M. A. D. & Plaxco, K. W. Apparent Debye-Huckel electrostatic effects in the folding of a simple, single domain protein. *Biochemistry* **44**, 1243–1250 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements T.L.R. was supported by a Trinity College Internal Graduate Studentship, and J.S.M. by a Gates Cambridge Scholarship. We would like to thank T. J. Rutherford and L. Kowalik for help with NMR, and J. M. Perez Canadillas for help with protein expression.

Author Contributions T.L.R. was responsible for the major experimental work and planning of this study, with further contributions from J.S.M., U.M. and S.M.V.F. A.R.F. wrote the paper with T.L.R. and performed data analysis.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.R.F. (arf25@cam.ac.uk).

RETRACTION

doi:10.1038/nature04183

Formation of zirconium metallic glass

Jianzhong Zhang & Yusheng Zhao

Nature 430, 332–335 (2004)

In this Letter, we concluded, primarily on the basis of energy-dispersive X-ray diffraction, that metallic glass could be formed in pure zirconium under high pressure and temperature conditions. However, careful observations using an angular-dispersive method and imaging-plate detector, together with X-ray-transparent anvils (H. Saitoh, T. Hattori, H. Kaneko, Y. Okajima and W. Utsumi, unpublished results), have revealed that our conclusion was in error. We are now convinced that the disappearance of diffraction lines of zirconium observed in our energy-dispersive experiments should be interpreted instead as rapid crystal growth at temperatures above that of the ω – β phase transformation.

Our original misinterpretation was partly due to our assumption that significant crystalline growth could not occur at temperatures of less than one-third of the melting point of zirconium. To a greater extent, it was a result of the limitations of the experimental techniques then available, which only allowed viewing of a very narrow window of Debye diffraction rings and therefore increased the probability of missing the Bragg spots. The state-of-the-art techniques used by Saitoh *et al.* at the Synchrotron Radiation Research Center, Japan Atomic Energy Institute, Hyogo, are superior for high-pressure research on non-crystalline materials.

We thank our Japanese colleagues for their careful work and for inviting us to observe their high pressure and temperature synchrotron X-ray diffraction experiments at beamline BL14B1 of SPring-8.

RETRACTION

doi:10.1038/nature04182

Negative lattice expansion from the superconductivity-antiferromagnetism crossover in ruthenium copper oxides

A. C. McLaughlin, F. Sher & J. P. Attfield

Nature 436, 829–832 (2005)

This Letter described several notable physical properties for low-doped ruthenium copper oxides, from which conclusions concerning superconductivity in copper oxides were drawn. A key result was the observation of an unusual, negative, expansion of the lattice parameters and volume at temperatures below the Ru spin-ordering transition, T_{Ru} , as shown in Figure 3. Unfortunately, we have discovered that a discrepancy in the algorithm for fitting the diffraction data led to erroneous shifts in the cell parameters and resulted in the apparent negative expansion. We retract the claim of negative

lattice expansion and the consequent inferences concerning the existence of two states in the pseudogap regime of the copper oxide electronic phase diagram. However, an anomalous expansion of the separation between copper oxide planes below T_{Ru} , as shown in Figure 3b, is reproducible and has subsequently been observed in other ruthenium copper oxide samples. Our results concerning large magnetoresistances and successive Ru and Cu spin-ordering transitions are not affected, so Figures 1 and 2 and the Supplementary Figures remain valid.

We thank S. Kimber and A. Williams for undertaking further calculations to determine the nature of the fitting problem.

RETRACTION

doi:10.1038/nature04181

RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription

Vera Schramke, Daniel M. Sheedy, Ahmet M. Denli, Carolina Bonila, Karl Ekwall, Gregory J. Hannon & Robin C. Allshire

Nature 435, 1275–1279 (2005)

An earlier paper¹ by two of us, in which it was concluded that expression of a hairpin RNA homologous to *ura4* RNA in *Schizosaccharomyces pombe* results in the production of siRNAs that bring about methylation of histone H3 on lysine 9 and recruitment of Swi6 to the *ura4* gene, has been retracted². This is because we are unable to reproduce the experiments involving expression of hairpin *ura4* RNA because the plasmid *sh-ura4SE-280*, or fission yeast strains that contain it, do not exist. Attempts to reproduce these observations with other *ura4*-hairpin constructs have failed.

As a result, we cannot verify the observations reported in our Letter to *Nature* and shown in Figures 1, 2 and 4a (lower panel) because these utilize the same plasmid *sh-ura4-280* hairpin construct. We therefore retract the conclusions that transcription of an siRNA homologous target is required to bring about silent chromatin assembly, that transcription of an siRNA homologous target by T7 is not sufficient to bring about silent chromatin assembly, and that transcription of an siRNA homologous target is required to allow association of Ago1 with that target.

However, we have confirmed that truncation of the RNAPol II CTD (removal of 17/28 heptad repeats) does alleviate silencing at centromeres, that centromeric siRNAs remain detectable (Figure 3), and that RNA pol II and Ago1 co-immunoprecipitate. In addition, the expression-profiling data generated for *rpb1-11* are robust.

V.S. has declined to sign this retraction as she maintains that the original conclusions are correct.

- Schramke, V. & Allshire, R. C. Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science* 301, 1069–1074 (2003).
- Schramke, V. & Allshire, R. C. Retraction. *Science* 310, 49 (2005).

-  FOCUS
-  SPOTLIGHT
-  RECRUITMENT
-  ANNOUNCEMENTS
-  EVENTS

naturejobs

The exception to the rule

For postdocs at the University of Texas Southwestern Medical Center in Dallas, the old adage 'be careful what you wish for, because it might come true' seems apt. Earlier this year, a survey by its postdoc association found that 84% of the centre's postdocs supported the idea of a campus-wide policy on issues such as salary.

Since then, the postdocs have been granted their wish: the campus has instituted a uniform pay policy. But the stipend levels are below the recommended minima set by the US National Institutes of Health (NIH), and well below salaries offered by industry or private research foundations.

Despite this, there is much good to say about both the concept and the process. Postdoc pay at many institutions varies — sometimes widely — from department to department, and even between people with similar training and experience. At Texas, the medical centre worked with postdoc representatives from each of its 13 departments to set more uniform pay standards, and succeeded in levelling the playing field.

But just how level should that playing field be? Although the idea of pay parity seems sensible, there is some

movement away from it, at least in special cases. Some organizations are offering what I call 'super postdocs'. These fellowships have much higher salaries and responsibilities and can serve as a much-needed bridge between a postdoc and a position as an independent investigator. Should those programmes be eliminated in the name of parity? Most sensible scientists would say no.

So how about a compromise, taking the best elements of all these approaches? Postdocs should certainly be involved in the stipend negotiation process. But rather than a ceiling, these stipend levels should represent a floor, that at least matches the lowest NIH recommended level. And they should allow exceptional compensation for exceptional cases. Such a system would be well worth wishing for.



Paul Smaglik, *Naturejobs* editor

CONTACTS

Publisher: Ben Crowe
Editor: Paul Smaglik
Marketing Manager: David Bowen

European Head Office, London
 The Macmillan Building, 4 Crinan Street
 London N1 9XW, UK
 Tel: +44 (0) 20 7843 4961
 Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com

Naturejobs Sales Director:
 Nevin Bayoumi (4978)
European Sales Manager:
 Andy Douglas (4975)

Natureevents: Sille Opstrup (4994)
UK/RoW/Ireland/Italy:
 Nils Moeller (4953)
 Irene Viglia-Atton (4944)
Scandinavia/Spain/Portugal:
 Evelina Rubio Håkansson (4973)
France/Switzerland/Belgium:
 Amelie Pequignot (4974)
Germany/Austria/The Netherlands:
 Reya Silao (4970)

Advertising Production Manager:
 Billie Franklin
 To send materials use London
 address above.
 Tel: +44 (0) 20 7843 4814

Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Niamh Shields

European Satellite Office
Germany/Austria/
The Netherlands:
 Patrick Phelan
 Tel: +49 89 54 90 57 11
 Fax: +49 89 54 90 57 20
 e-mail: p.phelan@nature.com

US Head Office, New York
 345 Park Avenue South,
 10th Floor, New York,

NY 10010-1707
 Tel: +1 800 989 7718
 Fax: +1 800 989 7103
 e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
 Chiyoda Building,
 2-37 Ichigayatamachi,
 Shinjuku-ku,
 Tokyo 162-0843
 Tel: +81 3 3267 8751
 Fax: +81 3 3267 8746
Asia-Pacific Sales Director: Rinoko Asami
 e-mail: rasami@naturejpn.com

PRIMED FOR A BIOTECH BOOM

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

One Osaka company has become the bellwether for Japanese biotechnology. AnGes, the first university spin-out to be listed on the Japanese stock exchange, is waiting for the results of a phase III clinical trial on a gene therapy to treat vascular ailments ranging from arteriosclerosis to angina. When they come — sometime next year — they could have a wide-ranging effect on Japan's nascent biotech industry.

"If it fails, investors might feel they shouldn't put money in biotech," says AnGes' co-founder Ryuichi Morishita, a gene-therapy expert at Osaka University. Investors and policy-makers alike hope that successful commercialization of the treatment could galvanize the nation's flagging biotechnology industry.

That the success of Japanese biotech might ride on the fortunes of an Osaka-based company will not surprise many in the country. Osaka gave rise to most of Japan's major pharmaceutical companies. The nation's drug industry has sales of some US\$60 billion and the pharmaceutical companies grouped in greater Osaka, centred in an area called Doshomachi, account for 13% of this total, more than any other region in the country. Moreover, Osaka University is arguably the country's strongest in life sciences, accounting for many of Japan's most influential biology papers over the past two decades.

Osaka's bid to become Japan's biomedical centre is buoyed by regional ties with nearby Kyoto University and the city of Kobe, where the country is focusing its efforts in developmental biology and regenerative medicine. Spring-8, the world's largest synchrotron, is only 120 kilometres away and easily accessible for

Biologists in Osaka think that their city's 'un-Japanese' culture makes it the ideal part of the country to become a hub for biotechnology. David Cyranoski investigates.

those who want to use the light source's radiation to analyse protein structures.

But Osaka's biotechnology ambitions are not immune from the obstacles faced by the rest of the country. Nationwide, industry and academia have shown little will to work together. Dynamic personalities with scientific, legal and business backgrounds who could bridge the gap between academia and industry remain few and far between — neither the culture nor the education system encourages people to dabble in different fields. Language difficulties and a reclusive culture mean there are few research links with international groups.

Confident approach

Nevertheless, the city is brimming with confidence and energy. In a country that often looks down on the profit motive, Osaka has traditionally been the merchant centre, where wholesalers gather and sell goods with entrepreneurial vigour — a stark contrast to the polished, bureaucratic and slow-moving Tokyo. "We have different DNA," says Morishita. Indeed, in the year following last spring's reorganization of universities into semi-private entities responsible for balancing their own books (see *Nature* **419**, 875–876; 2002), Osaka University chalked up more profit than any of Japan's 89 other former national universities. "We won't hesitate to make money," says molecular biologist Tsuneaki Sakata, a visiting professor at Osaka University. If a profitable biotechnology industry is going to flourish in Japan, Osaka may well be the place to look for the first shoots.

Osaka's biotechnology efforts are concentrated in the northern part of the city, which the central

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

government has designated a 'biocluster'. Between 2001 and 2004 the country invested ¥24 billion (US\$210 million) in the development of what its organizers at the Osaka Chamber of Commerce and Industry hope will rival successful US bioclusters, such as Genetown in Boston.

The 743-hectare Saito Life Science Park, opened in the spring of 2004, aims to lure bioventures and international drug companies with tax breaks and other incentives. The park also houses a national research institute for the development of new medicines. Some 20 biotechnology start-ups, including AnGes, are setting up shop there. "Things are really starting to pick up," says Sakata.

The park will build on the existing network of facilities and researchers, which includes the Osaka Bioscience Institute, Osaka University, the university hospital and the National Cardiovascular Center. The university and affiliated research institutes are a breeding-ground for scientific results with great potential. Between them these institutes have scored notable successes in areas such as protein structure and immunology.

Despite Osaka's wheeler-dealer reputation, most academic researchers there — like those throughout the country — continue to view involvement with industry as a distraction from their real work.

"Japan is losing a lot of its potential," says Toshikazu Nakamura, a geneticist at Osaka University. He learnt about patenting after his discovery in the 1980s of a protein called HGF, which stimulates the growth of many tissues, including blood cells. Unusually for a Japanese academic, Nakamura holds some 80 patents.



Toshio Yanagida is embracing Osaka's collaborative spirit to bring some cell biology into electrical engineering.

Academics' attitudes may be changing slowly since last year's reorganization, but so far the anticipated breakthrough in university-pharmaceutical cooperation has not taken place, says Toshio Tanaka, chief corporate adviser at drug company Tanabe Seiyaku in Osaka.

Commercial breaks

The Japanese drug industry remains sceptical about what academic researchers have to offer. The apparent mismatch stems from academic researchers' traditionally uncommercial outlook — even those who patent their discoveries are not necessarily doing the sort of cutting-edge research that companies want to invest in. "You can end up just piling up patents," says Tanaka. The situation is in stark contrast with the United States, where savvy academics know what will sell, he adds: "At US universities, when the professors put out something new, the businesses come running."

Another obstacle to bridging the biotechnology gap is the lack of multi-talented staff who can fill the gaps between academia, industry and law. Osaka hosts a leading biotech intellectual-property firm, Shusaku Yamamoto, named after its founder. On the day of *Nature's* visit to his office, Yamamoto had hired two biology PhDs — both from abroad. It is very difficult to find qualified scientists in Japan who have additional training in business or law, he says. Indeed, Yamamoto has opened a branch in Tokyo to broaden his search for capable multiskilled staff.

Those trying to make Osaka a biotechnology hub are addressing the problems of commercializing research head-on. The Osaka Chamber of Commerce

and Industry has established a 'biobusiness station,' a six-month training programme that invites speakers such as Yamamoto to teach scientists about the various fields of expertise, such as technology transfer, that are necessary in biotechnology. Since it was founded in February 2003, 98 graduates have passed through this training course and eight have taken up posts in industry. It is unusually diverse in the strict Japanese educational system, in which it is almost unheard-of for a biologist, say, to switch to law.

Osaka University, too, has been actively pushing industrial collaboration, says its president Hideo Miyahara. The university now has six formal collaborations with industry. The number of joint research projects with industry jumped from 206 to 508 in the year to March 2005.

National policies are supporting these efforts. Regulations precluding university researchers from working as board members for private enterprises have been relaxed. "Five years ago, the door was tightly closed," says Morishita. "Now it's quite a bit easier to work with industry."

Some now argue that the pressure to take up patentable research has gone too far. For example, a lot of researchers take exception to some grants that make applicants — even graduate students — list their patents and assess the applicability of research. "It's complex for researchers like me," says Akira Shinohara, a geneticist at Osaka University's Institute for Protein Research. "In the future, my work might be patentable, but it is difficult to know now."

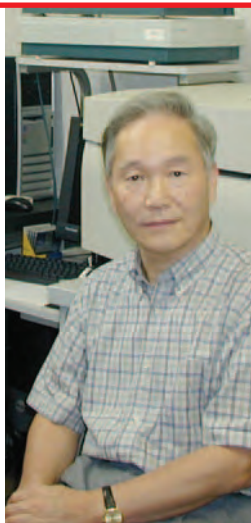
Fresh perspectives

But many in Osaka are embracing opportunities. Some interesting examples of the collaborative research between industry and academia are at the Graduate School of Frontier Biosciences, which focuses on interdisciplinary research. Toshio Yanagida, an engineer by trade, began monitoring the physical work of single-molecule proteins as they change the structure of the cell. He was impressed with the cell's ability to transfer information at nanoscale levels and high speed without producing destabilizing heat — something that designers working on semiconductors are finding increasingly hard to do.

In collaboration with electronics giant Matsushita, Yanagida is now transferring this knowledge of cell mechanics back to electrical engineering. He suggests that what has been called 'background noise' in a semiconductor system might actually be able to amplify the signal rather than drowning it out as would be expected. "It's time to take another look at noise," says Yanagida.

Such unconventional studies would be difficult elsewhere in Japan, says immunologist Tadamitsu Kishimoto, who established the graduate school during his term as university president, which ended in 2003. "There's more freedom in Osaka than in rigid, bureaucratic Tokyo," he says.

Other interdisciplinary research is already paying off. Soshio, a company at Osaka University's biotechnology incubator, struck it big when a team of electrical engineers and molecular biologists found that they could solve the structure of large membrane proteins by training a laser on them. Although such molecules account for a third of all drug targets, only a small percentage of these structures has so far been solved.



"Japan is losing a lot of its potential."
—Toshikazu Nakamura

The unprecedented efficiency in crystallizing proteins brought in an abundance of customers both in basic research and drug development. Within two months of opening its doors for business this July, Soshio was in the black.

"This could change the world of protein studies, and we don't have to go to the Moon to do it," says laser engineer Yusuke Mori, Soshio's representative director, referring to efforts to crystallize proteins in the space shuttle. He thanks the city's sense of adventure for opening up interdisciplinary studies. "In Osaka, many people want to take the challenge of exploring new fields," he says.

Breaking the barriers

Despite these successes, the Japanese working environment has not proved too appealing to foreigners. Language barriers and a bureaucratic system create obstacles to overseas companies and researchers.

Japanese companies likewise have done little to push beyond their borders. Drug companies depend on a large, and rather protected, domestic market rather than innovation, says Yamamoto. "If the focus is only on Japan, it's impossible to grow," he says.

The Institute for Protein Research has made international collaboration a priority by creating a full-time post for a foreign researcher. It also runs the Protein Data Bank of Japan, which shares structural information with the Worldwide Protein Data Bank. In August, the university's Research Institute for Microbial Diseases opened a centre in Thailand for collaborative research in infectious disease — a very rare move for a Japanese university.

Still, it is not clear how Osaka can clear the hurdles that have prevented it from becoming a magnet to business and investment like its regional rival Singapore, where multinational companies and internationally renowned researchers have set up shop.

Problems with the nationalized medicine policy make it difficult for companies, both domestic and foreign, to innovate, says Tanaka. He says that the health ministry forces the prices of drugs down even during the life of a patent, so that a new drug initially fetching ¥100 will raise only ¥20 by the end of its patent. "If there really are revolutionary ideas, we are very willing to invest. But when the prices keep falling, it's not worth it," Tanaka says.

Inadequate infrastructure for clinical trials, such as the absence of a payment for participants, makes it hard to gather patients, says Morishita. "In the United States you can do it for half the cost in a quarter of the time," he says. Distance from Tokyo makes wrangling over regulatory procedures even more difficult.

Nevertheless, an increasing number of companies are giving it a go. AnGes has just begun phase I trials on its second product, an oligonucleotide that intercepts immune-response signals in inflammatory diseases such as arthritis. But AnGes will have an earlier test in the shape of its first drug trial — and that could set the tone not only for the company's future but also for that of Osaka's biotech ambitions.

David Cyranoski is *Nature's* Asian-Pacific correspondent.

WEB LINKS

Osaka Bioscience Institute
 ▶ www.obi.or.jp/index2.html
 AnGes
 ▶ www.anges-mg.com/en/index.html
 Saito Life Science Park
 ▶ www.saito-kokubun.co.jp/eng/index.html
 Research Institute for Microbial Diseases
 ▶ www.biken.osaka-u.ac.jp/e
 Institute for Protein Research
 ▶ www.protein.osaka-u.ac.jp/home_e/index_e.html
 Graduate School of Frontier Biosciences
 ▶ www.fbs.osaka-u.ac.jp/eng/index.html

Toy planes

Time to leave.

Tobias S. Buckell

My sister Joanie's deft hands flicked from dreadlock to dreadlock, considering her strategy. "You always leaving," she said, flicking the razor on, and suddenly I'm five, chasing her with a kite made from plastic bags and twigs, shouting that I was going to fly away from her one day.

"I'm sorry. Please, let's get this done."

I'd waited long enough. I'd grown dreads because when I studied in the United States I wanted to remember who I was and where I came from as I began to lose my Caribbean accent. But the rocket plane's sponsor wanted them cut. It would be disaster for a helmet not to have a proper seal in an emergency. Explosive decompression was not something a soda company wanted to be associated with in their customers' minds. It was insulting that they assumed we couldn't keep the craft sealed. But we needed their money. The locks had become enough a part of me that I winced when the clippers bit into them, groaned, and another piece of me fell away.

In the back of the bus that I had pick me up, I hung on to a looped handle swinging from the roof as the driver rocketed down the dirt road from Joanie's. My sister had found a place out in the country, a nice concrete house with a basement opening up into a sloped garden on the side of a steep hill. She taught mathematics at the school a few miles away, an open-shuttered building, and this would have been my future too, if I hadn't been so intent on 'getting off the rock.'

The islands always called their children back.

We hit asphalt, potholes, and passed cane fields with machete-wielding labourers hacking away at the stalks, sweat-drenched shirts knotted around their waists. It was hot; my arms stuck to the plastic covered seats. The driver leaned into a turn, and looked back. "I want ask you something." I really wished the back seats had belts.

"Sure."

"All that money you spending, you don't think it better spent on getting better

roads?" He dodged a pothole. "Or more school funding?"

Colourful red and yellow houses on stilts dotted the steep lush green mountainsides as I looked out of the tinted windows. "Only one small part of the programme got funded by the government," I explained. "We found private investors, advertisers, to back the rest. Whatever the government invested will be repaid."

"Maybe."

I had my extra arguments. How many people lived on this island? Tens of thousands. Most of our food was imported, leav-

in front of it I was five years old again, with no money, and a piece of scrap metal in the triangular shape of a space plane. I would pretend it was just like the real life ones I'd read about in the books donated to the school after the hurricane. And at night, when the power would sometimes flicker out, I'd go out and stand on the porch and look up at the bright stars and envy them.

The stall had a small bottle, hammered over with soda-can metal, with triangular welded-on wings, and a cone stuck to the back. It was painted over in yellow, black and green, and I bought it.

The rest of the day was a blur. Getting to the field involved running the press: yes I'd cut my hair for 'safety' reasons, yes I thought this was a good use of our money, not just first-world nations deserved space, it was there for everyone.

There were photos of me getting aboard the tiny rocket plane with a small brown package under my arm. The giant balloon platform that the plane hung from shifted in the gentle, salty island breeze. Not too far away the waves hit the sand of the beach. Inside, suited up, door closed, everything became electronic.

It was the cheapest way to get to orbit. Balloon up on a triangular platform to save on fuel, then light the rocket-plane up and head for orbit. We'd scavenged balloons and material from several companies, one about to go out of business. The plane chassis had once been used by a Chinese corporation during trials, and the guidance systems were all open-source. Online betting parlours had our odds at 50%. We weren't even the first, but we were the first island.

The countdown finished, my stomach lurched, and I saw palm trees slide by the portholes to my right. I reached back and patted the package, the hammered-together toy, and smiled.

"Hello out there, all of you," I whispered into the radio. "We're coming up too." ■ Tobias S. Buckell is a Caribbean-born speculative fiction writer who now lives in Ohio. His first novel, *Crystal Rain*, will be out from Tor Books in February 2006. For more, see his website: www.TobiasBuckell.com.



JACEY

ing us dependent on other food-producing nations, who all used satellites to track their farming. What spin-off technologies might come out of studying recycling in space? Why wait for other nations to get to it first? Research always produced good things for the people who engaged in it.

But I was tired of arguing for it, and I had only sound bites for him, the same ones I'd given the media who treated us like kids trying to do something all grown up.

The market surrounded me in a riot of colour: fruit, vegetables, full women in dresses in bright floral patterns. And the noise of hundreds constantly bargaining over things like the price of fish. Teenagers stood around the corners with friends. I wandered around looking for something, as we needed to fill the craft with enough extra weight to simulate a passenger and we still had a few extra ounces to add.

I found a small toy stall. And standing